CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

210913Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 210-913
Supplement #:	Not applicable (Original-1)
Drug Name:	Cequa (OTX-101; cyclosporine ophthalmic solution) 0.09%
Indication(s):	To increase tear production (b) (4) associated with keratoconjunctivitis sicca (dry eye disease)
Applicant:	Sun Pharma Global FZE
Date(s):	Date Submitted: October 16, 2017
	Primary Review Due Date: July 12, 2018
	PDUFA Goal Date: August 16, 2018
Review Priority:	Standard
Biometrics Division:	DB IV
Statistical Reviewer:	Wonyul Lee, Ph.D.
Concurring Reviewers:	Yan Wang, Ph.D., Statistical Team Leader
Medical Division:	Division of Transplant and Ophthalmology Products
Clinical Team:	Lucious Lim, MD, Medical Officer
	William Boyd, MD, Medical Team Leader
Project Manager:	Lois Almoza

Keywords: double-masked, dry eye disease, generalized estimating equation, repeated measure ANCOVA

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

In this submission, the Applicant seeks approval of Cequa (OTX-101; cyclosporine ophthalmic solution) 0.09% to increase tear production (b) (4) associated with keratoconjunctivitis sicca (KCS), which is known as dry eye disease. The Applicant conducted two pivotal efficacy studies to evaluate OTX-101: OTX-101-20**14**-001 and OTX-101-20**16**-001. These studies are hereafter referred to as Study 14 and Study 16, respectively.

Studies 14 and 16 were multi-center, randomized, double-masked, vehicle-controlled, superiority studies. Study 14 randomized 455 subjects in a 1:1:1 ratio to OTX-101 0.05%, OTX-101 0.09%, or vehicle. In Study 16, a total of 745 subjects were randomized in a 1:1 ratio to OTX-101 0.09% or vehicle. In both studies, subjects were instructed to administer study medication topically to both eyes twice daily for 12 weeks. The primary efficacy evaluations were conducted at Day 84.

Study 14 had co-primary endpoints of sign and symptom. The sign primary endpoint was the mean change from baseline in total conjunctival staining score. The symptom primary endpoint was the mean change from baseline in global symptom score. On the other hand, Study 16 had a single primary endpoint associated with tear production: response rate in the Schirmer's test score defined as the proportion of eyes achieving ≥ 10 mm increase from baseline in the Schirmer's test score. The selection of the primary endpoint in Study 16 was based on the positive results of the Schirmer's test score in Study 14.

For the filing of this NDA submission, the Applicant proposed to use the response rate in the Schirmer's test score as a single primary endpoint. This proposal was considered acceptable by the Division of Transplant and Ophthalmology Products (DTOP) at the End-of-Phase 2 meeting. The DTOP also agreed that the analysis results of this endpoint for Study 14 can be considered supportive. Note that the approval of RESTASIS® in 2003 was based on the same endpoint for the indication of increasing tear production in a certain subpopulation of KCS.

In both studies, the response rate in the Schirmer's test score was significantly higher in the OTX-101 0.09% group compared to that in the vehicle group: 16.8% vs. 8.6% in Study 14 and 16.6% vs. 9.2% in Study 16 (Table 1). The differences in the response rates (OTX-101 - vehicle) were statistically significant: 8.2% [95% CI: (1.9%, 14.6%)] in Study 14 and 7.3% [95% CI: (3.3%, 11.3%)] in Study 16.

Table 1: Proportion of eyes with ≥10 mm increase from baseline in the Schirmer's test score at Day 84

	OTX-101-2014-001		**OTX-101-2016-001**	
	OTX-101 0.09% N = 152	**Vehicle N = 152**	**OTX-101 0.09% N = 371**	**Vehicle N = 373**
Number of eyes	304	304	742	746
Eyes with ≥10-mm increase, n (%) [1]	51 (16.8%)	26 (8.6%)	123 (16.6%)	69 (9.2%)
Difference [95% CI] [2]	8.2% [1.9%, 14.6%]		7.3% [3.3%, 11.3%]	
P-value for Difference [2]	0.0113		0.0003	

[1] All randomized and treated subjects were included in the analysis. Missing values were treated as failures;
[2] Based on a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes of the same subject were treated as repeated measures; CI: confidence interval.

The two pivotal studies also evaluated the efficacy of OTX-101 in terms of the ocular surface integrity. The efficacy variables associated with the ocular surface integrity included (1) total lissamine green conjunctival staining score (0-12 scale) and (2) total fluorescein corneal staining score (0-20 scale). See Section 3.2.1 for details regarding the definitions of these total scores. In principle, lower total staining score represents better integrity of the ocular surface.

(b) (4)

Regarding safety, the OTX-101 0.09% group in Study 16 had higher rate of adverse events (AE) compared to the vehicle group: 40.9% vs. 26.3%. The most common AE in the OTX-101 0.09% group was instillation site pain. Excluding instillation site pain, however, the AE profile of OTX-101 0.09% was not significantly different from that of vehicle: 27.2% vs. 23.7%. In Study 14, the two treatment groups had comparable AE rates: 34.9% vs. 34.2%. The reviewer defers to the medical reviews for a comprehensive safety evaluation.

In summary, the reviewer concludes that this application provided adequate statistical evidence of efficacy to support an approval of OTX-101 0.09% ophthalmic solution for the proposed indication of increasing tear production (as measured by the Schirmer's test score) (b) (4) associated with KCS.

2 INTRODUCTION

This section provides an overview of the application, a summary of the clinical studies selected for review, and information on data sources for review.

2.1 Overview

The Applicant seeks approval of OTX-101ophthalmic solution (Cequa) 0.09% to increase tear production (b) (4) associated with keratoconjunctivitis sicca (KCS), which is also referred to as dry eye disease.

KCS, one of the most common ocular conditions, is a complex disease of tears and ocular surface. While its etiology is not fully understood, it is recognized that triggers of KCS are multifactorial including autoimmune disorders, certain medications, eye surgeries, and environmental stresses. Associated eye symptoms include dryness, irritation, stinging, burning, redness, pain, fatigue, and blurred vision.

FDA approved two drugs for the treatment of KCS: RESTASIS® (cyclosporine ophthalmic emulsion) 0.05% and XIIDRA® (lifitegrast ophthalmic solution) 5%. RESTASIS® is indicated to increase tear production in patients whose tear production is presumed to be suppressed due to ocular inflammation associated with KCS. On the other hand, XIIDRA® is indicated for the treatment of the signs and symptoms of KCS.

2.1.1 Class and Indication

Per the Applicant, OTX-101 ophthalmic solution is a clear, preservative-free, and nanomicellar formulation of cyclosporine. The indication that the Applicant seeks is to increase tear production (b) (4) associated with KCS.

2.1.2 History of Drug Development

Ocular Technologies SARL has conducted and sponsored clinical studies for the development of OTX-101 ophthalmic solution under IND 118954. Then, the current applicant, Sun Pharma Global FZE, acquired Ocular Technologies SARL in January 2017 with all the rights to OTX-101 ophthalmic solution.

The sponsor, Ocular Technologies SARL, discussed the clinical development program with the DTOP at appropriate development milestones. A pre-IND meeting was on October 23, 2013. The sponsor filed the initial IND application on July 23, 2014 to conduct a Phase 2/3 study (Study 14). This study had co-primary efficacy endpoints of sign and symptom (b) (4) However, Study 14 failed to show a significant treatment effect on the primary symptom endpoint. Consequently, based on the results of Study 14, the Sponsor designed a Phase 3 study (Study 16) that had only one primary sign endpoint associated with tear production. The protocol of Study 16 was submitted for the End-of-Phase 2 meeting held on October 6, 2015 and the DTOP agreed with the proposed primary endpoint as indicated in the following comments to the sponsor:

- *"We would accept a statistically significant difference in the proportion of subjects with ≥ 10 mm increase in a Schirmer's test score from baseline as a single primary efficacy endpoint for NDA filing, provided that NDA is supported by adequate safety and efficacy data. Approval of the NDA is a review issue."*
- *"Study OTX-101-2014-001 could be considered supportive data. An additional adequate and well controlled study demonstrating an increase in tear production is highly recommended prior to submission of an NDA."*

Following completion of Study 16, the Applicant had a type B pre-NDA meeting with the DTOP on April 24, 2017.

Reviewer's note: The reviewer also finds this single primary efficacy endpoint acceptable considering that the approval of RESTASIS® in 2003 was based on the same endpoint. In the clinical studies of RESTASIS®, the response rate in the Schirmer wetting was approximately 15% for the RESTASIS®-treated groups and 5% for the vehicle-treated groups.

2.1.3 Specific Studies Reviewed

The sponsor conducted the following four clinical studies under IND 118954:

- **Study OTX-101-2014-001**: Phase 2/3 dose-ranging study
- **Study OTX-101-2016-001**: Phase 3 efficacy and safety study
- Study OTX-101-2016-002: Ongoing safety extension of the Phase 3 study
- Study OTX-101-2016-003: Phase 1 PK study

This review focuses on Studies OTX-101-20**14**-001 and OTX-101-20**16**-001 as the efficacy and safety of Cequa were primarily evaluated in these studies. The two studies are hereafter referred to as Study 14 and Study 16, respectively. A summary of the two studies is presented in Table 3.

Table 3: Summary of specific studies reviewed

	OTX-101-2014-001 (Phase 2/3)	**OTX-101-2016-001** (Phase 3)
Design	A multi-center, randomized, double-masked, vehicle-controlled, superiority study	
Duration	12-week treatment (84 days)	
Treatment / Sample Size	OTX-101 0.09% BID/ 152 OTX-101 0.05% BID / 151 Vehicle BID / 152	OTX-101 0.09% BID / 372 Vehicle BID / 373
Primary Endpoint	Co-Primary: - Sign: Mean change from baseline at Day 84 in total conjunctival staining score in the study eye - Symptom: Mean change from baseline at Day 84 in global symptom score	Single Primary: Proportion of eyes with ≥ 10 mm increase from baseline in the Schirmer's test score at Day 84 based on data for both eyes
Key Secondary Endpoints	- Mean change from baseline at Day 84 in the Schirmer's test score - Mean change from baseline at Day 84 in total corneal staining score	- Mean change from baseline at Day 84 in total conjunctival staining score - Mean change from baseline at Day 84 in central corneal staining score
Study Population	- Age 18 years or older with clinical diagnosis of bilateral KCS - Patient-reported history of KCS for a period of at least 6 months	

2.2 Data Sources

The data sources for this review included protocols, statistical analysis plans (SAPs), clinical study reports (CSRs), the summary of clinical efficacy (SCE), the summary of clinical safety (SCS), and the datasets for the respective studies. The datasets were submitted in the formats of Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) in electronic submission. The datasets can be located at \\CDSESUB1\evsprod\NDA210913\0000\m5\datasets and \\CDSESUB1\evsprod\NDA210913\0004\m5\datasets. The first location also includes the SAS programs used to generate tables and figures for the efficacy and safety analyses in the clinical study reports. The second location includes only the analysis dataset of lissamine green conjunctival staining score for Study 16 that was accidently excluded in the first location.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

No major issues were identified regarding the quality and integrity of the submitted SDTM and ADaM datasets. The data quality control/assurance procedures were properly documented in the clinical study reports.

The datasets were well organized. The Applicant's primary efficacy results were reproducible using the ADaM datasets. In general, using SDTM and ADaM datasets, the reviewer was able to conduct the necessary analyses without complex manipulations.

3.2 Evaluation of Efficacy

This section evaluates the efficacy results of Studies 14 and 16.

3.2.1 Study Design and Endpoints

Study Design

Studies 14 and 16 had similar designs. They were randomized, multi-center, double-masked, vehicle-controlled, superiority studies comparing OTX-101 to vehicle for the treatment of KCS. Randomization was as follows:

- Study 14 randomized 455 subjects in a 1:1:1 ratio to OTX-101 0.05%, OTX-101 0.09%, or vehicle. Randomization was stratified by study site. A total of 29 sites enrolled subjects in the United States.
- Study 16 randomized 745 subjects in a 1:1 ratio to OTX-101 0.09% or vehicle. Randomization was not stratified by study site. A total of 45 sites enrolled subjects in the United States.

The study schedule was as follows (see Appendix G for more details):

- Screening/run-in period: Subjects who met the study criteria entered a run-in period for 14 days. In the run-in period, subjects administered vehicle topically twice daily (BID) to both eyes. The run-in period could be extended up to 17 days in Study 14 and 20 days in Study 16.

- Baseline (Day 0): Eligible subjects were randomized. Baseline ophthalmic examinations were conducted. Subjects administered the first dose of study medication in a supervised manner.
- Following the Baseline visit, subjects administered study medication topically to both eyes BID for 12 weeks (Day 1 to Day 84).
- Safety and efficacy evaluations were conducted at the following visits:
  - Study 14: Days 14 (±3), 28 (±3), 42 (±3), 56 (±3), and 84 (+7)
  - Study 16: Days 28 (±3), 56 (±3), and 84 (+7)

The key inclusion criteria were as follows:

- Age 18 years or older
- Patient-reported history of KCS for a period of at least 6 months
- Clinical diagnosis of bilateral KCS
- Lissamine green conjunctival staining sum score of ≥ 3 to ≤ 9 out of a total possible score of 12 (scoring excluded superior zones 2 and 4) in the same eye at Screening and Baseline.
- Global symptom score ≥ 40 at Screening and Baseline
- Visual acuity (VA)
  - Study 14: Snellen visual acuity of better than 20/200 in each eye
  - Study 16: Corrected Snellen visual acuity of better than 20/200 in each eye

Study 14 designated the study eye as the eye with the higher conjunctival staining score at baseline. If both eyes had the same score, the right eye was the study eye. On the other hand, Study 16 did not designate study eye.

Primary Efficacy Endpoints

Study 14 had co-primary efficacy endpoints of sign and symptom as follows:

- Sign: the mean change from baseline at Day 84 in total conjunctival staining score in the study eye;
- Symptom: the mean change from baseline at Day 84 in global symptom score.

The initial goal of the program was to demonstrate superiority of OTX-101 to vehicle for both the sign and symptom endpoints. However, as Study 14 failed to meet the symptom primary endpoint, the sponsor proposed to use the following single primary efficacy endpoint for the NDA filing in this submission (see Section 2.1.2):

- **Proportion of eyes demonstrating ≥ 10 mm increase from baseline in the Schirmer's test score at Day 84**.

For concise presentation, the proportion of eyes demonstrating ≥ 10 mm increase from baseline is referred to as "response rate" in this review. Basically, the Schirmer's test measures the amount of tears produced. During the Schirmer's test, strips are placed in both eyes at the same time for 5 minutes. After 5 minutes, the amount of wetting in the strips is recorded in mm.

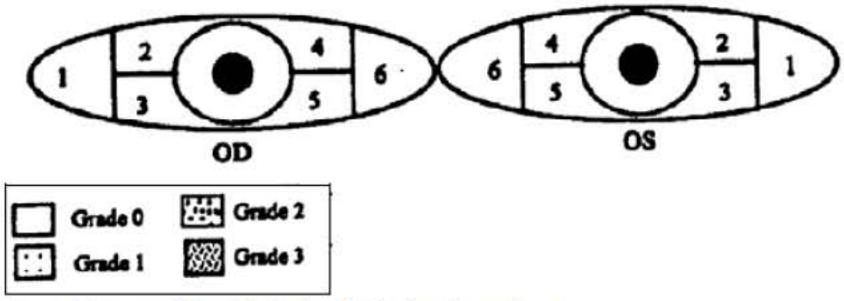
Efficacy Measures Associated With Ocular Surface Signs

Recall that the sponsor's proposed indication is to increase tear production and [redacted] (b) (4) associated with KCS. Thus, in addition to the response rate in the Schirmer's test score, this review also focuses on evaluating efficacy measures associated with ocular surface signs.

Staining is one of the most common methods to quantify ocular surface integrity. The two pivotal studies used lissamine green conjunctival staining and fluorescein corneal staining to measure the integrity of conjunctival surface and corneal surface, respectively.

In the lissamine green conjunctival staining, each of six areas of conjunctiva was graded in 0-3 scale as shown in the figure below. In principle, lower score represents better integrity of the conjunctival surface. For each eye, the Applicant defined the total conjunctival staining score as the sum of the scores over zones 1 (temporal), 3, 5 (inferior), and 6 (nasal), excluding zones 2 and 4 (superior). Thus, the total conjunctival staining score for each eye ranges from 0 to 12.

Figure 2: Division of Conjunctival Surface for Grading



Source: Figure 2 of the statistical analysis plan for Study 16.

In the fluorescein corneal staining, each of five areas of cornea was scored in 0-4 scale with 0.5 increments as shown in the figure in the next page. In principle, lower score represents better integrity of the corneal surface. For each eye, the Applicant defined the total corneal staining score as the sum of the scores over the all five areas: central, superior, inferior, lateral, and medial. Thus, the total corneal staining score for each eye ranges from 0 to 20.

Efficacy Endpoints To Be Focused In This Review

This review focuses on [redacted] (b) (4) the proposed indication of increasing tear production [redacted] (b) (4) :

- **The response rate in the Schirmer's test score at Day 84;**
- **Mean change from baseline in total conjunctival staining score at Day 84;**
- **Mean change from baseline in total corneal staining score at Day 84.**

[redacted] (b) (4) See Appendix A for the full list of the secondary and additional endpoints in the two pivotal studies.

<u>Reviewer's note</u>: The mean change from baseline in total conjunctival staining score was one of the co-primary endpoints in Study 14 and one of the secondary endpoints in Study 16. The mean change from baseline in total corneal staining score was one of the secondary endpoints in Study 14 and an additional endpoint in Study 16.

Figure 1: Division of Corneal Surface for Grading

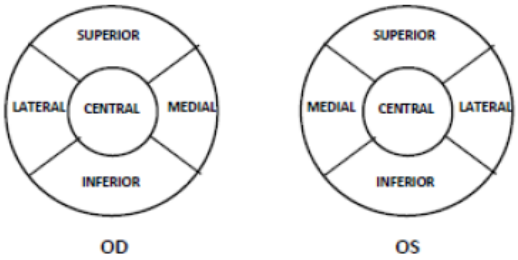


Table 1: Expanded NEI/Industry Workshop Scale

Score	Description
0	No punctate stain in area
0.5	1-5 micropunctate stain spots
1.0	6-15 micropunctate stain spots
1.5	More than 15 micropunctate stain spots
2.0	Moderate macropunctate stain spots (involving less than 50% of the area)
2.5	Moderate macropunctate stain spots (involving more than 50% of the area)
3.0	Clumped macropunctate stain spots (involving less than 50% of the area)
3.5	Clumped macropunctate stain spots (involving more than 50% of the area)
4.0	Severe diffuse (coalescent) macropunctate stain of the area

Source: Figure 1 and Table 1 of the statistical analysis plan for Study 16.

Efficacy Measures Associated With Symptoms

Note that the proposed indications by the Applicant do not include any symptom indication. However, this review also evaluates symptom measures collected in the studies to have a more complete picture of the efficacy of OTX-101 0.09%.

The two pivotal studies used a modified SANDE instrument (Schaumberg et al, 2007) to assess dry eye symptoms. More specifically, subjects were asked the following two questions:

- Frequency: please indicate how often, over the past week, your eyes felt dry and/or irritated (0-100 scale; 0 = rarely; 100 = all the time);
- Severity: please indicate how severe, on average you felt your symptoms of dryness and/or irritation were over the past week (0-100 scale; 0 = very mild; 100 = very severe).

The global symptom score was defined as the square-root of the product of the two scores. The mean change from baseline at Day 84 in the global symptom score was one of the co-primary endpoints in Study 14 and one of the secondary endpoints in Study 16.

Study 16 had an additional symptom measure: Ocular Surface Disease Index (OSDI). OSDI consists of 12 questions. Each question was assessed in a 0-4 scale (See Appendix H for details). Lower score represents better symptom. The mean change from baseline at Day 84 in the visual-related function subscale (sum over questions 6-9 in OSDI) was one of the secondary endpoints.

3.2.2 Statistical Methodologies

This section primarily focuses on describing statistical methodologies for analyzing the three efficacy measures: the Schirmer's test score, lissamine green conjunctival staining score, and fluorescein corneal staining score. Statistical methodologies for analyzing symptom measures are also briefly described.

Analysis populations

The protocols and the statistical analysis plans defined three analysis populations as follows:

1. The intent-to-treat (ITT) population was defined as all randomized subjects. The primary efficacy analyses were performed on the ITT population.
2. The per-protocol (PP) population was defined as all ITT subjects who remained in the study through Day 84 and who did not deviate from the protocol in any way likely to seriously affect the primary outcome of the study. Supportive efficacy analyses were conducted on the PP population.
3. The safety population was defined as all randomized subjects who received at least one dose of the study medication. All safety analyses were performed on the safety population.

Reviewer's note: In Study 16, the ITT population excluded one subject (ID: (b) (6)) who was randomized to the OTX-101 0.09% group. This subject was never treated with the study medication.

Analysis methods for the efficacy endpoints

For both studies, the Applicant analyzed the response rate in the Schirmer's test score using a generalized estimating equation (GEE) model. The model included only treatment as a fixed effect. The reviewer found the following issues regarding the Applicant's analysis:

- The GEE model was supposed to treat responses from the left and right eyes of the same subject as correlated (repeated) measures. However, the Applicant's actual analysis treated them as independent. This was caused by a misspecification in their SAS programming codes (See Appendix B for details).
- The GEE model specified in the final SAP was supposed to use the logit link function for the binary response (increase of $\geq$ 10 mm). However, the Applicant's actual analysis used the identity link function instead of the logit link function.

As a result, the Applicant's actual analysis became equivalent to a simple two-sample Z-test ignoring the correlation between the left and right eyes. The reviewer observed that the Applicant's analysis results were exactly the same as the results of two-sample Z-tests.

The Applicant changed the link function (from the logit to the identity) to provide a 95% confidence interval for the difference in the response rates. We requested this confidence interval at the pre-NDA meeting as follows:

"Please provide a 95%confidence interval for the response rate for each treatment group and a 95% confidence interval for the difference in the response rates between the treatment groups. Please provide details on the methods for calculating these 95% confidence intervals."

To investigate validity of the results from the GEE model using the identity link, the reviewer conducted the following two analyses (See Appendix C for more details):

- The GEE model using the identity link was compared with a Bootstrap method requiring no parametric assumptions. In this analysis, the results from the GEE model were almost identical to the results from the Bootstrap method when the GEE model accounted for the correlation between the two eyes.
- The reviewer performed simulation studies to investigate coverage probabilities of 95% confidence intervals from the GEE model using the identity link. The simulated coverage probability was between 94% and 95% when the GEE model accounted for the correlation between the two eyes. On the other hand, the coverage probability was between 88% and 90% when the GEE model ignores the correlation as the Applicant's analysis did.

Based on these observations, the reviewer finds the GEE model using the identity link acceptable for this particular application as long as the GEE model accounts for the correlation between the left and right eyes of the same subject.

Therefore, in this review, the reviewer used the GEE model with the identity link to compute a 95% confidence interval and p-value for the difference in the response rates while accounting for the correlation between the left and right eyes of the same subject.

Recall that this review also focuses on the evaluation of the following two efficacy endpoints associated with ocular surface sign:

- Mean change from baseline in total conjunctival staining score at Day 84;
- Mean change from baseline in total corneal staining score at Day 84.

The primary analysis methods for these endpoints in the final SAPs were as follows:

- Study 14: an ANCOVA model including treatment, baseline value, and site.
- Study 16: a repeated measure ANCOVA model including treatment, baseline value, visit, and treatment by visit interaction. Per the SAP, observations from left and right eye were supposed to be treated as repeated measures.

In this NDA submission, the Applicant reanalyzed the data from Study 14 using the repeated measure ANCOVA model in Study 16 to allow direct comparison between the studies. The efficacy results in the Applicant's proposed labeling also present the results based on the repeated measure ANCOVA model for both studies. For Study 14, Section 3.2.4 (results and conclusions) of this review contains both (1) the original result from the ANCOVA model and (2) the reanalyzed result from the repeated measure ANCOVA model.

Reviewer's note: The Applicant mis-specified a certain part of the SAS programming codes for fitting the repeated measure ANCOVA model. As a result, the Applicant's actual analysis treated the observations from left and right eyes as independent rather than as repeated measures. The reviewer made proper corrections to the SAS programming codes to perform the analysis as planned in the final SAP.

Recall that this review also evaluates the following two symptom endpoints:

- Mean change from baseline at Day 84 in the global symptom score (SANDE);
- Mean change from baseline at Day 84 in the visual-related function subscale (Items 6-9) of the OSDI

These symptom endpoints were analyzed as specified in the SAPs as follows:

- Study 14: an ANCOVA model including treatment, baseline value, and site.
- Study 16: a repeated measure ANCOVA model including treatment, baseline value, visit, and treatment by visit interaction.

Handling of missing values

For the binary outcome of ≥ 10 mm increase from baseline in the Schirmer's test score, any missing value was treated as failure. For the continuous outcomes, missing values were not explicitly imputed for the repeated measure ANCOVA models. For the ANCOVA models in Study 14, missing values were imputed by the last observation carried forward (LOCF).

3.2.3 Subject Disposition, Demographic, and Baseline Characteristics

Subject disposition and primary reasons for study discontinuation are summarized in Table 4. In both studies, the proportion of the subjects who completed the studies was slightly lower in the OTX-101 0.09% group compared to that in the vehicle group: 92.1% vs. 94.7% in Study 14 and 93.3% vs. 96.8% in Study 16.

In Study 14, 29 subjects (6.4%) discontinued the study due to adverse events (6), withdrawal of subject consent (12), lost to follow-up (6), and other reasons (5). The adverse events were instillation site pain and conjunctival hyperaemia in the OTX-101 0.05% group, back pain in the OTX-101 0.09% group, and seasonal allergy, conjunctivitis viral, and cerebrovascular accident in the vehicle group. No death was reported.

In Study 16, 37 (5%) subjects discontinued the study due to adverse events (16), death (1), withdrawal of subject consent (9), lost to follow-up (8), protocol deviation (2), and investigator decision (1). The number of discontinued subjects due to adverse events was higher in the OTX-101 0.09% group compared to that in the vehicle group: 14 (3.8%) vs. 2 (0.5%). The most common adverse event leading to study discontinuation was the instillation site pain (9 in the OTX-101 0.09% and 0 in the vehicle). Other ocular adverse events leading to study discontinuation included eye irritation, eye pain, conjunctival hyperaemia, macular fibrosis, blurred vision, instillation site pruritus, and instillation site reaction. See Table 29 for the full list of the AEs that resulted in study discontinuation. The Applicant reported that the death in the OTX-101 0.09% group was considered unrelated to the study medication.

Table 4: Subject disposition, n (%)

	OTX-101-2014-001				OTX-101-2016-001		
	All	**OTX-101 0.05%**	**OTX-101 0.09%**	**Vehicle**	**All**	**OTX-101 0.09%**	**Vehicle**
Randomized Population	N = 455	N = 151	N = 152	N = 152	N = 745	N = 372	N = 373
Subjects Completed Study							
Yes	426 (93.6)	142 (94.0)	140 (92.1)	144 (94.7)	708 (95.0)	347 (93.3)	361 (96.8)
No	29 (6.4)	9 (6.0)	12 (7.9)	8 (5.3)	37 (5.0)	25 (6.7)	12 (3.2)
Reasons for Discontinuation							
Adverse Event	6 (1.3)	2 (1.3)	1 (0.7)	3 (2.0)	16 (2.1)	14 (3.8)	2 (0.5)
Death	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.1)	1 (0.3) †	0 (0.0)
Withdrawal of subject consent	12 (2.6)	3 (2.0)	7 (4.6)	2 (1.3)	9 (1.2)	5 (1.3)	4 (1.1)
Lost to Follow-up	6 (1.3)	1 (0.7)	2 (1.3)	3 (2.0)	8 (1.1)	4 (1.1)	4 (1.1)
Protocol Deviation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.3)	0 (0.0)	2 (0.5)
Investigator Decision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.1)	1 (0.3)	0 (0.0)
Other	5 (1.1)	3 (2.0)	2 (1.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

† 56-year old white female. Unknown causes. Reported not related to the study medication.
Source: Table 14.1.2.1 of the CSR for Study 14 and Table 6 of the CSR for Study 16.

Baseline demographic characteristics are summarized in Table 5. In general, the demographic characteristics were balanced between the OTX-101 0.09% group and the vehicle group. In terms of age, gender, and ethnicity, the reviewer did not identify any notable imbalance between the OTX-101 0.09% group and the vehicle group. With respect to race, in Study 14, the proportion of Black or African American was slightly lower in the OTX-101 0.09% group compared to that in the vehicle group: 8.6% vs. 13.2%.

The mean age was 60.2 and 59.0 years in Study 14 and Study 16, respectively. The majority of the subjects were female (79.3% in Study 14 and 84.2% in Study 16). In terms of race, most subjects were White (82.2% in Study 14 and 82.7% in Study 16) followed by Black or African American (9.7% in Study 14 and 11.5% in Study 16). Most subjects identified themselves as neither Hispanic nor Latino (82.6% in Study 14 and 85.1% in Study 16).

Table 5: Baseline demographic

	OTX-101-2014-001				OTX-101-2016-001		
	All	**OTX-101 0.05%**	**OTX-101 0.09%**	**Vehicle**	**All**	**OTX-101 0.09%**	**Vehicle**
Randomized Population	N = 455	N = 151	N = 152	N = 152	N = 745	N = 372	N = 373
Age							
Mean (SD)	60.2 (13.9)	62.0 (13.3)	59.2 (14.5)	59.4 (13.8)	59.0 (14.4)	58.4 (14.1)	59.5 (14.7)
Median	62.0	63.0	60.0	60.5	61.0	60.0	62.0
Min - Max	22 - 91	23 - 91	23 - 85	22 - 89	18 - 90	18 - 89	20 - 90
Age Category, n (%)							
< 65	261 (57.4)	81 (53.6)	89 (58.6)	91 (59.9)	455 (61.1)	240 (64.5)	215 (57.6)
>= 65	194 (42.6)	70 (46.4)	63 (41.4)	61 (40.1)	290 (38.9)	132 (35.5)	158 (42.4)
Gender, n (%)							
Female	361 (79.3)	119 (78.8)	122 (80.3)	120 (78.9)	627 (84.2)	316 (84.9)	311 (83.4)
Male	94 (20.7)	32 (21.2)	30 (19.7)	32 (21.1)	118 (15.8)	56 (15.1)	62 (16.6)

	OTX-101-2014-001				OTX-101-2016-001		
	All	OTX-101 0.05%	OTX-101 0.09%	Vehicle	All	OTX-101 0.09%	Vehicle
Ethnicity, n (%)							
Not Hispanic or Latino	376 (82.6)	131 (86.8)	123 (80.9)	122 (80.3)	634 (85.1)	315 (84.7)	319 (85.5)
Hispanic or Latino	79 (17.4)	20 (13.2)	29 (19.1)	30 (19.7)	111 (14.9)	57 (15.3)	54 (14.5)
Race, n (%)							
White	374 (82.2)	128 (84.8)	126 (82.9)	120 (78.9)	616 (82.7)	311 (83.6)	305 (81.8)
Black†	44 (9.7)	11 (7.3)	13 (8.6)	20 (13.2)	86 (11.5)	41 (11.0)	45 (12.1)
Asian	18 (4.0)	8 (5.3)	7 (4.6)	3 (2.0)	23 (3.1)	11 (3.0)	12 (3.2)
Other	16 (3.5)	3 (2.0)	6 (3.9)	7 (4.6)	18 (2.4)	8 (2.2)	10 (2.7)
American Indian	1 (0.2)	1 (0.7)	0 (0.0)	0 (0.0)	1 (0.1)	1 (0.3)	0 (0.0)
Native Hawaiian††	2 (0.4)	0 (0.0)	0 (0.0)	2 (1.3)	1 (0.1)	0 (0.0)	1 (0.3)

† Black or African American; †† Native Hawaiian Pacific Islander.
Source: Table 7 of the CSR for Study 14 and Table 8 of the CSR for Study 16.

Baseline sign and symptom scores are summarized in Table 6. This summary was produced using the average score of the left and right eyes. In general, the baseline sign and symptom scores were comparable between the OTX-101 0.09% group and the vehicle group in both studies.

Table 6: Baseline sign and symptom (average of both eyes)

	OTX-101-2014-001			OTX-101-2016-001	
	OTX-101 0.05%	OTX-101 0.09%	Vehicle	OTX-101 0.09%	Vehicle
ITT population					
Number of subjects	N = 151	N = 152	N = 152	N = 371	N = 373
Schirmer's Test Score					
Mean (SD)	10.73 (7.57)	12.34 (8.67)	12.19 (8.25)	11.89 (7.77)	12.09 (7.73)
Median	9.00	9.00	9.50	10.00	10.00
Min - Max	0.50 - 35.00	0.50 - 35.00	0.50 - 35.00	0.50 - 35.00	0.50 - 35.00
Total Conjunctival Staining Score					
Mean (SD)	5.72 (1.76)	5.40 (1.76)	5.52 (1.66)	5.42 (1.71)	5.52 (1.77)
Median	5.50	5.00	5.00	5.50	5.00
Min - Max	2.50 - 10.50	2.50 - 9.50	2.00 - 9.00	2.00 - 10.00	1.50 - 9.50
Total Corneal Staining Score					
Mean (SD)	5.00 (2.95)	4.40 (2.85)	4.42 (2.64)	4.06 (2.37)	4.30 (2.65)
Median	4.50	4.25	4.25	3.75	4.00
Min - Max	0.25 - 14.00	0.00 - 13.00	0.00 - 12.25	0.00 - 12.75	0.00 - 14.00
SANDE† Global Symptom Score					
Mean (SD)	61.77 (15.32)	62.62 (14.56)	61.49 (14.34)	63.06 (15.71)	62.24 (16.12)
Median	60.00	62.65	59.58	60.00	60.00
Min - Max	27.39 - 99.50	40.00 - 95.00	40.00 - 98.99	40.00 - 100.00	40.00 - 100.00

†SANDE: Symptom frequency and severity rating.
Source: Table 8 of the CSR for Study 14 and Table 9 of the CSR for Study 16.

Recall that Study 14 designated a study eye for each subject. See Table 22 in Appendix D for the baseline summaries by the study eyes and fellow eyes. No notable imbalance was observed between the OTX-101 0.09% group and the vehicle group. Compared to the fellow eyes, the study eyes had worse baseline outcomes (lower tear production and higher staining scores).

3.2.4 Results and Conclusions

Sections 3.2.4.1 – 3.2.4.3 provides efficacy results for [redacted] (b) (4): Schirmer's test score, [redacted] (b) (4). Section 3.2.4.4 provides efficacy results for the symptom measures: SANDE and OSDI. The reviewer's efficacy conclusion is provided in Section 3.2.4.5.

3.2.4.1 Schirmer's Test Score at Day 84

The two pivotal studies measured tear production for each eye by the Schirmer's test at baseline and Day 84. Higher test score represents more tear production. The primary efficacy endpoint in Study 16 was the response rate (proportion of eyes with ≥ 10 mm increase from baseline) in the Schirmer's test score at Day 84.

Table 7 shows the primary analysis results of this endpoint for the ITT population. In Study 16, the OTX-101 0.09% group showed higher response rate compared to the vehicle group: 16.6% vs. 9.2%. The treatment difference (OTX-101 0.09% - vehicle) was statistically significant: 7.3% [95% CI: (3.3%, 11.3%)]. Study 14 also showed higher response rate in the OTX-101 0.09% group compared to that in the vehicle group: 16.8% vs. 8.6%. The treatment difference was 8.2% [95% CI: (1.9%, 14.6%)].

Table 7: Proportion of eyes with ≥10 mm increase from baseline in the Schirmer's test score at Day 84 (ITT population, missing values were treated as failures)

	OTX-101-2014-001		OTX-101-2016-001	
	OTX-101 0.09%	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Number of eyes	304	304	742	746
Number of eyes with ≥10 mm increase, n (%) [1]	51 (16.8%)	26 (8.6%)	123 (16.6%)	69 (9.2%)
Difference in proportions (OTX-101 - Vehicle)		8.2%		7.3%
95% CI for Difference and P-value				
a. Applicant's analysis [2]	[3.0%, 13.5%], 0.0021		[3.9%, 10.7%], < 0.0001	
b. Reviewer's analysis [3]	[1.9%, 14.6%], 0.0113		[3.3%, 11.3%], 0.0003	

[1] All randomized and treated subjects were included in the computation. Missing values were treated as failures.
[2] The Applicant's analysis was equivalent to a simple two-sample Z-test ignoring correlation between left and right eyes.
[3] Based on a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes were treated as repeated measures.
Source: Table 6 of the SCE.

Reviewer's note: The Applicant's analysis resulted in narrower CIs and smaller p-values compared to those from the reviewer's analysis. This is because the Applicant's analysis did not account for the correlation between the left and right eyes of the same subject while the reviewer's analysis did. Note that the subjects administered the same drug to both eyes. In this setting, without accounting for the correlation between the two eyes, the standard errors for the

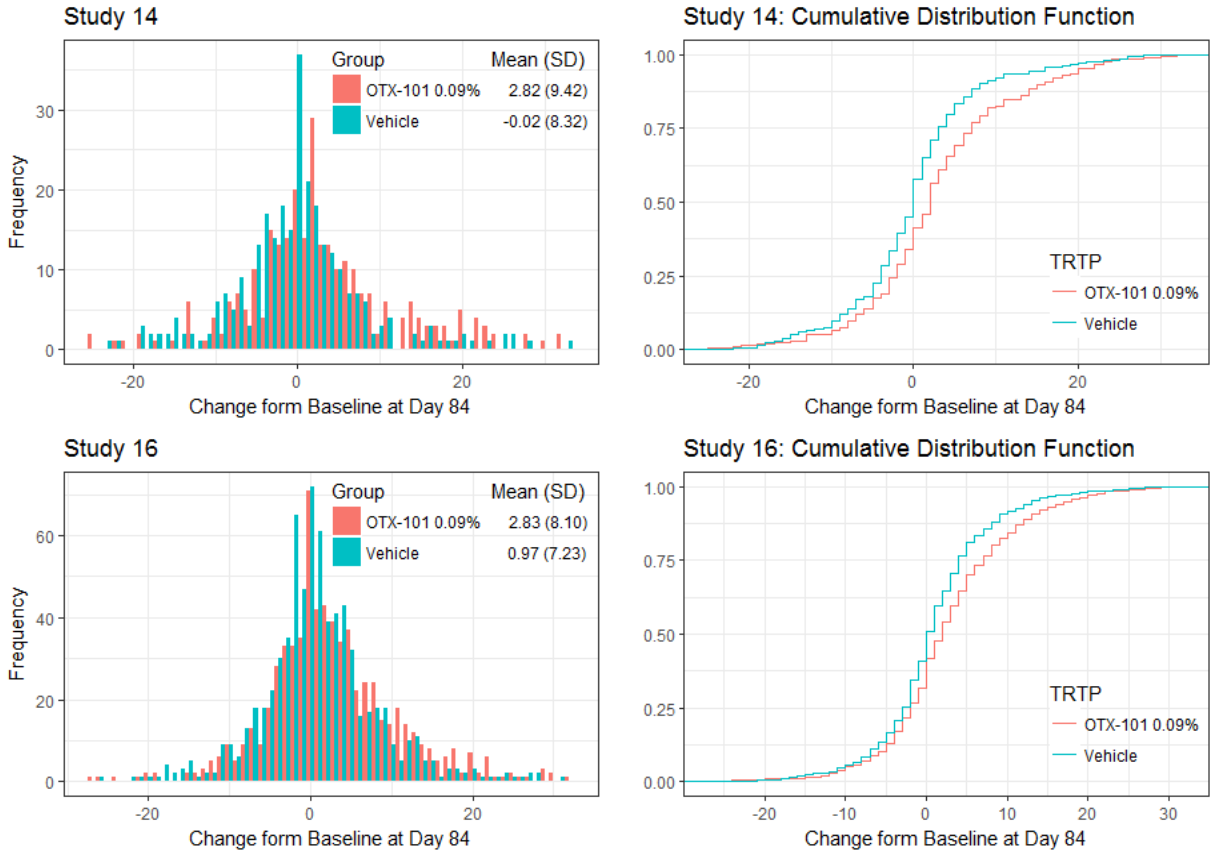
estimated treatment differences can be underestimated, which can result in invalid CIs and p-values. The reviewer's simulation study demonstrated this issue (See Appendix C).

A supportive analysis on the PP population resulted in the similar conclusion (See Table 23 in Appendix E). In this supportive analysis, the observed treatment difference was 9.6% [95% CI: (2.5%, 16.6%)] in Study 14 and 8.1% [95% CI: (3.9%, 12.5%)] in Study 16, respectively.

Another supportive analysis was performed using the average score of the left and right eyes (See Table 24 in Appendix E). The estimated treatment difference in this analysis was 9.2% [95% CI: (2.0%, 16.4%)] in Study 14 and 6.0% [95% CI: (1.7%, 10.3%)] in Study 16, respectively.

Figure 1 depicts distributions of the change from baseline at Day 84 in the Schirmer's test score. In Study 14, the mean change from baseline at Day 84 was 2.82 and -0.02 in the OTX-101 0.09% group and the vehicle group, respectively. Study 16 also shows higher mean change from baseline in the OTX-101 0.09% group compared to that in the vehicle group: 2.83 vs. 0.97. The treatment difference (OTX-101 0.09% - vehicle) was 2.84 [95% CI: (1.37, 4.30)] in Study 14 and 1.86 [95% CI: (1.05, 2.66)] in Study 16. The cumulative distribution functions (the right panels) show separation between the OTX-101 0.09% group and the vehicle group, which implying larger tear increase from baseline for the OTX-101 0.09% group.

Figure 1: Distribution of change from baseline at Day 84 in the Schirmer's test score (ITT population)



* Plots were produced using observed cases only without any imputation for missing values.

3.2.4.4 Global Symptom Score and OSDI

Table 12 presents the analysis results for the change from baseline in the global symptom score at Day 84. In both studies, the two treatment groups had similar level of reduction from baseline in the global symptom score: -19.16 vs. -19.90 in Study 14 and -18.54 vs. -19.09 in Study 16. The observed treatment differences were not significant.

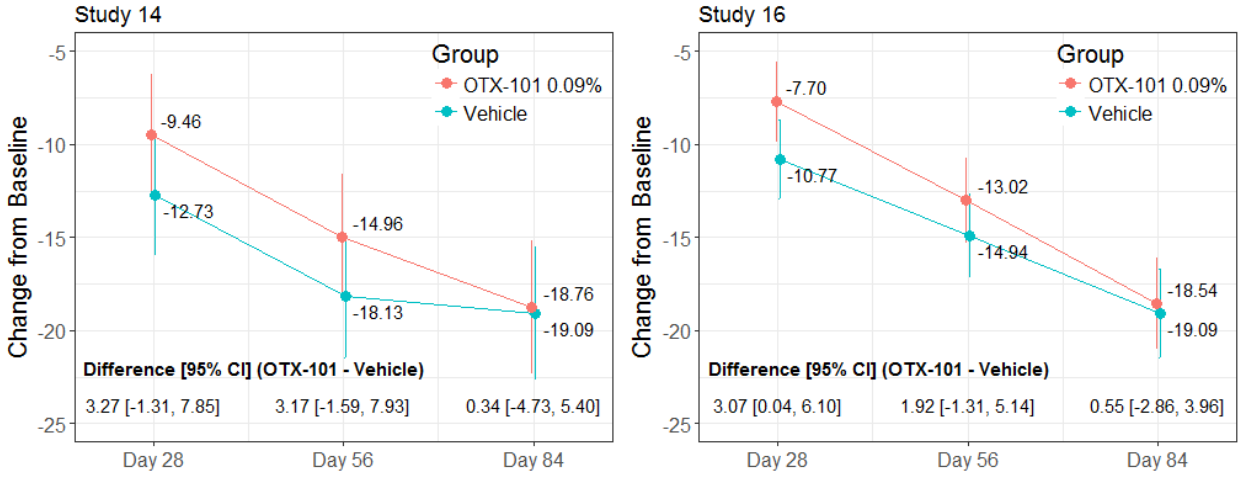
Table 12: Mean change from baseline in global symptom score at Day 84 (ITT population)

	OTX-101-2014-001[1]		OTX-101-2016-001[2]	
	OTX-101 0.09%	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Number of subjects	152	152	371	373
Mean change from baseline (SE)	-19.16 (1.92)	- 19.90 (1.90)	-18.54 (1.24)	-19.09 (1.22)
Difference [95% CI]		0.74 [-4.41, 5.89]		0.55 [-2.86, 3.96]
P-value for Difference		0.7783		0.7528

[1] Based on ANOCVA model including baseline value, treatment, and site.
[2] Based on repeated measure ANOCVA models including baseline value, treatment, visit, and treatment by visit interaction.
SE: standard error; CI: confidence interval.
Source: Table 11 of the CSR for Study 14 and Table 17 of the CSR for Study 16.

Figure 6 depicts the mean change from baseline in global symptom score at Days 28, 56, and 84. The vehicle group showed larger decrease in the global symptom score at Days 28 and 26 compared with the OTX-101 0.09% group. However, the decrease from baseline at Day 84 was similar between the two groups.

Figure 6: Mean change from baseline in the global symptom score over time (ITT population)



* Means, Differences, and 95% CIs were obtained using repeated measure ANCOVA models including baseline score, treatment, visit, and treatment by visit interaction.

Table 13 summarizes the analysis results for the change from baseline at Day 84 in OSDI for Study 16. The two treatment groups showed similar level of reduction from baseline: -3.93 vs. -4.04 in terms of total OSDI and -1.30 vs. -1.44 in terms of the visual-related function subscale of OSDI. The treatment differences were not significant.

[redacted] (b) (4)

3.2.4.5 Efficacy Conclusion

The two pivotal studies demonstrated evidence of efficacy of OTX-101 0.09% for the proposed indication of increasing tear production [redacted] (b) (4). In terms of tear production as measured by the Schirmer's test score, the response rate (increase of ≥ 10 mm from baseline) at Day 84 was significantly higher in the OTX-101 0.09% group compared to that of the vehicle group. The mean change from baseline in the Schirmer's score at Day 84 was also higher for the OTX-101 0.09% group. [redacted] (b) (4)

3.3 Evaluation of Safety

In this section, high-level summaries of adverse events are provided; see the FDA medical reviews for a comprehensive safety evaluation.

3.3.1 Extent of Exposure

Table 14 presents a summary of the exposure to the study medication. The subjects were instructed to administer study medication topically to both eyes for 12 weeks (Day 1 to Day 84). Mean duration of the exposure was comparable between the OTX-101 0.09% group and vehicle group: 11.6 vs. 12.0 in Study 14 and 11.9 vs. 12.3 in Study 16. No notable difference was identified.

Table 14: Exposure to study medication

	OTX-101-2014-001			OTX-101-2016-001	
	OTX-101 0.05%	**OTX-101 0.09%**	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Safety Population	N = 151	N = 152	N = 152	N = 372	N = 372
Exposure (weeks)					
Mean (SD)	11.8 (2.2)	11.6 (2.4)	12.0 (1.8)	11.9 (2.4)	12.3 (1.0)
Median	12.1	12.1	12.1	12.1	12.1
Min - Max	0.6 - 13.3	1.1 - 13.4	1.0 - 14.1	0.1 - 15.3	3.1 - 15.1

Source: Table 2 of the Summary of Clinical Safety.

3.3.2 Summary of Adverse Events

Table 15 presents a high-level summary of adverse events (AE). In Study 14, the proportion of subjects with AEs was comparable between the OTX-101 0.09% group and vehicle group. However, in Study 16, the OTX-101 0.09% group showed higher proportion of subjects with AEs compared to the vehicle group: 40.9% vs. 26.3%. **The most common AE in the OTX-101 0.09% group was instillation site pain**. It appears that the imbalance in the AE rate between the two treatment groups was mostly driven by instillation site pain. The reviewer observed that the proportion of subjects with AEs excluding instillation site pain was comparable between the two groups: 27.2% vs. 23.7%.

Table 15: Summary of adverse events (AEs)

	OTX-101-2014-001			OTX-101-2016-001	
	OTX-101 0.05%	**OTX-101 0.09%**	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Safety Population	151	152	152	372	372
Subjects with AE	47 (31.1%)	53 (34.9%)	52 (34.2%)	152 (40.9%)	98 (26.3%)
Subjects with Serious AE	1 (0.7%)	0 (0%)	4 (2.6%)	6 (1.6%)	2 (0.5%)
Subjects withdrawn due to AE	4 (2.6%)	5 (3.3%)	6 (3.9%)	17 (4.6%)	3 (0.8%)
Subjects with TEAE	44 (29.1%)	52 (34.2%)	51 (33.6%)	151 (40.6%)	91 (24.5%)
- Mild TEAE	25 (16.6%)	30 (19.7%)	26 (17.1%)	111 (29.8%)	74 (19.9%)
- Moderate TEAE	14 (9.3%)	19 (12.5%)	23 (15.1%)	34 (9.1%)	15 (4%)
- Severe TEAE	5 (3.3%)	3 (2%)	2 (1.3%)	6 (1.6%)	2 (0.5%)

* AE: Adverse Event, TEAE: Treatment-Emergent Adverse Event.
Source: Table 7 of the Summary of Clinical Safety.

Regarding study discontinuation due to AEs, 15 subjects in Study 14 and 20 subjects in Study 16 discontinued the studies due to AEs. The AEs leading to the study discontinuation for at least 2 subjects are presented in Table 16. In the table, the number in the parenthesis is the number of AEs suspected to be related to the study medication. **The most common AE leading to study discontinuation was instillation site pain** (3 in Study 14 and 9 in Study 16) followed by eye irritation and eye pain. The full list of AEs leading to study discontinuation can be found in Table 29 of Appendix F.

Table 16: AEs leading to study discontinuation for at least 2 subjects

	OTX-101-2014-001			OTX-101-2016-001	
	OTX-101 0.05%	**OTX-101 0.09%**	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Subjects withdrawn due to AE	4 (4)	5 (3)	6 (1)	17 (12)	3 (2)
Preferred Term					
Instillation site pain	2 (2)	1 (1)	0 (0)	9 (8)	0 (0)
Eye irritation	0 (0)	1 (1)	0 (0)	1 (1)	0 (0)
Eye pain	0 (0)	0 (0)	0 (0)	2 (2)	0 (0)

* The number in the parenthesis is the number of AEs suspected to be related to the study medication.
Source: Table 14.3.1.5 in the CSR for each of Study 14 and Study 16.

In terms of serious AE, a total of 5 subjects in Study 14 and 8 subjects in Study 16 experienced serious AEs. Table 17 presents the serious AEs for them. None of them were reported as related to the study medication.

Table 17: Serious AEs

	OTX-101-2014-001			OTX-101-2016-001	
	OTX-101 0.05%	**OTX-101 0.09%**	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Subjects with Serious AEs	1	0	4	6	2
Serious AE					
Bipolar disorder	0	0	1	0	0
Cerebrovascular accident	0	0	1	0	0
Death	0	0	0	1	0
Humerus fracture	0	0	1	0	0
Hypokalaemia	0	0	1	0	0
Lung neoplasm malignant	0	0	0	1	0
Nephrolithiasis	0	0	0	1	0
Perforated ulcer	0	0	0	0	1
Peripheral nerve decompression	1	0	0	0	0
Pneumonia	0	0	0	1	0
Spinal column stenosis	0	0	0	1	0
Spinal osteoarthritis	0	0	0	0	1
Subdural haematoma	0	0	0	1	0

* None of the serious AEs were reported as related to the study medication.
Source: Table 28 of the CSR for Study 14 and Table 25 of the CSR for Study 16.

In terms of treatment-emergent adverse event (TEAE), the OTX-101 0.09% group and vehicle group in Study 14 had similar proportions of subjects with TEAE. In Study 16, the OTX-101 0.09% group showed higher proportion of subjects with TEAE compared to the vehicle group: 40.6% vs. 24.5%. In both studies, the majority of TEAEs were mild or moderate. **The most common ocular TEAE in the OTX-101 0.09% group was instillation site pain** (See Table 18). By excluding instillation site pain, the TEAE rates were comparable between the two groups: 26.1% in the OTX-101 0.09% vs. 22% in the vehicle.

3.3.3 Safety Conclusion

In Study 14, the AE profile of the OTX-101 0.09% was comparable to that of the vehicle. However, in Study 16, the OTX-101 0.09% group showed higher AE rate compared to the vehicle group. The imbalance in the AE rates was mostly driven by the instillation site pain, which was the most common AE in the OTX-101 0.09% group. No remarkable difference between the treatment groups was observed for the other AE categories. In both studies, the majority of AEs were either mild or moderate. In summary, the AE profile of the OTX-101 0.09% appeared similar to that of the vehicle except instillation site pain; however, deference is made to the FDA medical reviews for a comprehensive safety evaluation and conclusion.

Table 18: Most common TEAEs

	OTX-101-2014-001				OTX-101-2016-001			
	OTX-101 0.09%		Vehicle		OTX-101 0.09%		Vehicle	
	Events	Subjects	Events	Subjects	Events	Subjects	Events	Subjects
Safety Population	152		152		372		372	
Subjects with TEAE	52 (34.2%)		51 (33.6%)		151 (40.6%)		91 (24.5%)	
Ocular TEAE	48	37 (24.3%)	40	28 (18.4%)	186	125 (33.6%)	91	66 (17.7%)
Instillation site pain	25	24 (15.8%)	5	5 (3.3%)	95	90 (24.2%)	17	16 (4.3%)
Conjunctival hyperaemia	0	0 (0.0%)	0	0 (0.0%)	33	30 (8.1%)	24	19 (5.1%)
Eye irritation	3	3 (2.0%)	1	1 (0.7%)	3	3 (0.8%)	6	5 (1.3%)
Eye pruritus	1	1 (0.7%)	3	3 (2.0%)	1	1 (0.3%)	5	5 (1.3%)
Foreign body sensation in eyes	1	1 (0.7%)	0	0 (0.0%)	1	1 (0.3%)	5	5 (1.3%)
Non-ocular TEAE	28	21 (13.8%)	44	32 (21.1%)	81	49 (13.2%)	45	33 (8.9%)
Urinary tract infection	2	2 (1.3%)	2	2 (1.3%)	4	4 (1.1%)	2	2 (0.5%)
Headache	2	2 (1.3%)	0	0 (0.0%)	9	6 (1.6%)	2	2 (0.5%)
Sinusitis	0	0 (0.0%)	0	0 (0.0%)	4	4 (1.1%)	5	5 (1.3%)
Upper respiratory tract infection	3	3 (2.0%)	3	3 (2.0%)	1	1 (0.3%)	2	2 (0.5%)
Bronchitis	2	2 (1.3%)	2	2 (1.3%)	3	3 (0.8%)	1	1 (0.3%)

Source: Table 14.3.1.2 in the CSR for each of Study 14 and Study 16.

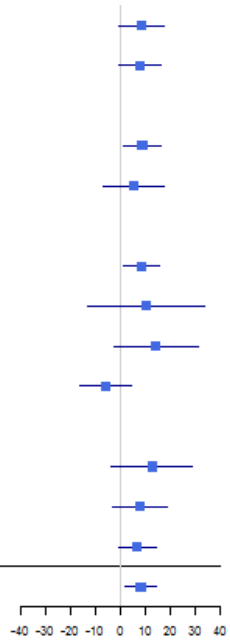
4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The response rate in the Schirmer's test score at Day 84 was analyzed across subgroups defined by age (< 65 years or ≥ 65 years), gender, race, or baseline score (< 5 mm, 5-10 mm, or ≥ 10 mm). For each subgroup, a 95% confidence interval for the difference in the response rates was obtained using the GEE model described in Section 3.2.2.

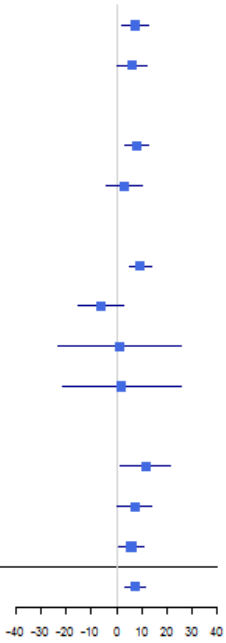
Table 19 presents the results of the subgroup analyses. In Study 14, the observed treatment difference was consistently positive across all subgroups except the "Other" race group. Note that the sample size in the "Other" subgroup is small (12 in the OTX-101 and 18 in the vehicle). Thus, the reviewer does not consider this negative effect as a notable subgroup finding. In Study 16, the observed treatment difference in the Black subgroup was negative in favor of the vehicle group. Considering that this negative effect was not observed in Study 14, the reviewer cannot make any confirmative conclusion of this negative result in the Black subgroup. In the subgroup analysis by the baseline score, the observed treatment difference was consistently positive across the subgroups in favor of the OTX-101 0.09%.

Table 19: Proportion of eyes with ≥10 mm increase from baseline in the Schirmer's test score at Day 84 by age, gender, race, or baseline score subgroups (ITT population, missing values were treated as failures)

Study 14	OTX-101 0.09%	Vehicle	Difference
Age			
< 65	36/178 (20.2%)	21/182 (11.5%)	8.7% [-0.3, 17.7]
>= 65	15/126 (11.9%)	5/122 (4.1%)	7.8% [-0.5, 16.1]
Gender			
Female	44/244 (18.0%)	22/240 (9.2%)	8.9% [1.5, 16.2]
Male	7/60 (11.7%)	4/64 (6.2%)	5.4% [-6.9, 17.7]
Race			
White	43/252 (17.1%)	20/240 (8.3%)	8.7% [1.6, 15.8]
Black	6/26 (23.1%)	5/40 (12.5%)	10.6% [-12.9, 34.0]
Asian	2/14 (14.3%)	0/6 (0.0%)	14.3% [-2.4, 31.0]
Other	0/12 (0.0%)	1/18 (5.6%)	-5.6% [-15.8, 4.7]
Baseline Score			
< 5 mm	10/48 (20.8%)	4/45 (8.9%)	12.7% [-3.5, 28.9]
5-10 mm	21/112 (18.8%)	12/111 (10.8%)	7.9% [-2.9, 18.7]
>= 10 mm	20/144 (13.9%)	10/148 (6.8%)	6.9% [-0.6, 14.5]
All Subjects	51/304 (16.8%)	26/304 (8.6%)	8.2% [1.9, 14.6]



Study 16	OTX-101 0.09%	Vehicle	Difference
Age			
< 65	87/478 (18.2%)	46/430 (10.7%)	7.5% [2.2, 12.8]
>= 65	36/264 (13.6%)	23/316 (7.3%)	6.4% [0.4, 12.4]
Gender			
Female	113/630 (17.9%)	62/622 (10.0%)	8.0% [3.5, 12.5]
Male	10/112 (8.9%)	7/124 (5.6%)	3.3% [-3.9, 10.5]
Race			
White	112/620 (18.1%)	52/610 (8.5%)	9.5% [5.1, 14.0]
Black	6/82 (7.3%)	12/90 (13.3%)	-6.0% [-15.0, 3.0]
Asian	3/22 (13.6%)	3/24 (12.5%)	1.1% [-23.7, 25.9]
Other	2/18 (11.1%)	2/22 (9.1%)	2.0% [-21.5, 25.5]
Baseline Score			
< 5 mm	28/137 (20.4%)	12/118 (10.2%)	11.6% [1.5, 21.7]
5-10 mm	39/226 (17.3%)	23/227 (10.1%)	7.4% [0.5, 14.2]
>= 10 mm	56/379 (14.8%)	34/401 (8.5%)	5.9% [1.0, 10.9]
All Subjects	123/742 (16.6%)	69/746 (9.2%)	7.3% [3.3, 11.3]



* The 95% confidence intervals for differences (OTX-101 0.09% – vehicle) were obtained using a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes were treated as repeated measures.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The reviewer did not identify major statistical issues that can impact the overall conclusions. Minor issues identified during the review were as follows:

(1) Analysis of the response rate (proportion of subjects with ≥ 10 mm increase from baseline in the Schirmer's test score)

- Per the primary analysis method (the GEE model) specified in the final SAP, binary responses from left and right eyes were supposed to be treated as repeated measures to account for correlation between them. However, in the actual Applicant's analysis, the two eyes were treated as independent due to misspecification in a certain part of the SAS programming codes.
- In this review, the reviewer used the GEE model while accounting for the correlation between the two eyes. The reviewer's analysis resulted in slightly wider 95% CIs than the Applicant's analysis which did not account for the correlation. The reviewer's results were also almost identical to the results from the Bootstrap method (See Appendix C).

(b) (4)

5.2 Collective Evidence

The Applicant seeks approval of OTX-101 0.09% to increase tear production (b) (4) associated with KCS. Efficacy and safety were evaluated in the two pivotal studies, Study 14 and Study 16.

In terms of tear production as measured by the Schirmer's test, the two pivotal studies showed that the proportion of eyes with ≥ 10 mm increase from baseline at Day 84 was significantly higher in the OTX-101 0.09% group compared to that of the vehicle group: 16.8% vs. 8.6% in Study 14 and 16.6% vs. 9.2% in Study 16, respectively. The difference in the response rates (OTX-101 - vehicle) was statistically significant: 8.2% [95% CI: (1.9%, 14.6%); p-value of 0.0113] in Study 14 and 7.3% [95% CI: (3.3%, 11.3%); p-value of 0.0003] in Study 16. The mean change from the baseline at Day 84 was also larger in the OTX-101 group compared to the vehicle group: 2.82 vs. -0.02 in Study 14 and 2.83 vs. 0.97 in Study 16. The treatment difference was 2.84 [95% CI: (1.37, 4.30)] in Study 14 and 1.86 [95% CI: (1.05, 2.66)] in Study 16.

(b) (4)

(b) (4)

The AE profile of OTX-101 0.09% was comparable to that of vehicle in Study 14. However, in Study 16, OTX-101 0.09% group had higher rate of AEs compared to the vehicle group: 40.9% vs. 26.3%. The most common AE was instillation site pain. Excluding instillation site pain, the AE profile of OTX-101 0.09% was comparable to that of vehicle: 27.2% vs. 23.7%.

5.3 Conclusions and Recommendations

OTX-101 (cyclosporine ophthalmic solution) 0.09% was shown to be superior to the vehicle in increasing tear production (b) (4). The statistical analysis results provided adequate evidence of efficacy to support an approval of OTX-101 0.09% ophthalmic solution for the proposed indication.

5.4 Labeling Recommendations

Section 14 (Clinical Studies) in the Applicant's proposal presents the efficacy results of tear production (b) (4) for the two pivotal studies. As the Applicant's analyses did not account for the correlation between left and right eyes, the precision of their analysis results can be slightly misleading. In addition, the Applicant's proposed labeling lacks some details regarding the analysis methods. Thus, we recommend the following tables presenting the reviewer's analysis results and some details of the analysis methods:

Table 1: Tear Production

	Study 1		**Study 2**	
	Cequa N = 152	**Vehicle N = 152**	**Cequa N = 371**	**Vehicle N = 373**
Number of eyes	304	304	742	746
Eyes with ≥ 10-mm increase in tear production at Day 84, n (%) [1]	51 (16.8%)	26 (8.6%)	123 (16.6%)	69 (9.2%)
Difference (Cequa - Vehicle) [95% CI] [2]	8.2% [1.9%, 14.6%]		7.3% [3.3%, 11.3%]	

[1] Eyes of all randomized and treated subjects were included in the analysis. Missing values were treated as failures to demonstrating an increase of ≥ 10-mm increase.
[2] Based on a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes were treated as repeated measures. P-value for the difference was < 0.02 in Study 1 and < 0.001 in Study 2.


(b) (4)

APPENDICES

Appendix A. Secondary and Additional Efficacy Endpoints in Two Pivotal Studies

OTX-101-2014-001	OTX-101-2016-001
Secondary Efficacy Endpoints	
1. Mean change from baseline in TBUT in the study eye from baseline to Day 84.	1. Mean change from baseline in lissamine green conjunctival staining score (excluding superior zones) at Visit 5 (Day 84) based on data for both eyes.
2. Mean change from baseline in total corneal fluorescein staining score in the study eye at Day 84.	2. Mean change from baseline in Schirmer's test (unanesthetized) at Visit 5 (Day 84) based on data for both eyes.
3. Patient satisfaction with treatment score (5-point ordinal scale) at Day 84.	3. Mean change from baseline in central corneal staining score at Visit 5 (Day 84) based on data for both eyes.
4. Mean change from baseline in Schirmer's test score (categorized) at Day 84 (average of both eyes).	4. Complete clearing of central corneal fluorescein staining at Visit 5 (Day 84) based on data for both eyes.
	5. Mean change from baseline in SANDE global symptom score at Visit 5 (Day 84).
	6. Mean change from baseline in the visual-related function subscale (Items 6 – 9) of the OSDI questionnaire at Visit 5 (Day 84).
Additional Efficacy Endpoints	
1. Proportion of subjects demonstrating a ≥ 30% reduction in total conjunctival staining score in the study eye from baseline to Day 84.	1. Complete clearing of lissamine green conjunctival staining in the temporal zone at Visit 5 (Day 84) based on data for both eyes.
2. Mean change from baseline at Days 14, 28, 42 and 56 in the study eye for: a. Total conjunctival staining score b. Global symptom score c. TBUT d. Total corneal fluorescein staining score	2. Repeated measures analysis of all time points for: a. Lissamine green conjunctival staining score (excluding superior zones), b. Total fluorescein corneal staining score, c. SANDE global symptom score.
	3. Mean change from baseline at Visit 5 (Day 84) for the total OSDI score and for each of its subscales.
3. Patient satisfaction with treatment score (5-point scale) at Days 28 and 56	4. The frequency and intensity components of the SANDE global symptom score at Visit 5 (Day 84).

SANDE = Symptom Assessment iN Dry Eye; TBUT = tear break-up time
Source: Table 1 of the Summary of Clinical Efficacy.

Appendix B. Issue with the Applicant's SAS programming Codes

The reviewer noticed a misspecification in the Applicant's SAS programming codes for the GEE model. (b) (4) This resulted in treating the two eyes from a subject as independent. The followings are the Applicant's SAS programming codes:

(b) (4)

Here, the (b) (4) categorical variable indicating left eye or right eye. (b) (4) With the code above, the reviewer observed that the resulting GEE correlation matrix for repeated measures had dimension of 1, which means that the two eyes were not considered as repeated measures.

Appendix C. Additional Analysis and Simulation Study to Assess the GEE Model

C-1: Comparison with a Bootstrap method

The reviewer applied a Bootstrap method to compute a 95% confidence interval (CI) for the difference in the response rates in the Schirmer's test score. The steps for the Bootstrap method were as follows:

- Step 1: **Randomly draw n subjects from the ITT population with replacement**. Here, n is the number of subjects in the ITT population. The resulting sample of size n is referred to as a Bootstrap sample.
- Step 2: Generate 10,000 Bootstrap samples.
- Step 3: For each of 10,000 Bootstrap samples, compute the difference in the response rates. The resulting 10,000 values constitute a Bootstrap distribution of the difference in the response rates.
- Step 4: The final 95% CI is [L, U] where L is the lower 2.5% percentile and the upper 97.5% percentile of the Bootstrap distribution.

As the Bootstrap sampling was performed at the subject level, the correlation between the left and right eyes of the same subject was naturally accounted for. Note that the Bootstrap method does not require any distributional assumptions.

Table 20 shows the Bootstrap 95% CIs and p-values for the difference in the response rates in the Schirmer's test score. The 95% CI from the Bootstrap method was [1.8%, 14.6%] in Study 14 and [3.4%, 11.3%] in Study 16. These results were almost identical to the results from the GEE model with the identity link function: [1.9%, 14.6%] in Study 14 and [3.3%, 11.3%] in Study 16.

Table 20: Proportion of eyes with ≥10 mm increase from baseline in the Schirmer's test score at Day 84 (ITT population, missing values were treated as failures).

	OTX-101-2014-001		OTX-101-2016-001	
	OTX-101 0.09%	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Number of eyes	304	304	742	746
Number of eyes with ≥10 mm increase, n (%) [1]	51 (16.8%)	26 (8.6%)	123 (16.6%)	69 (9.2%)
Difference in proportions (OTX-101 - Vehicle)		8.2%		7.3%
95% CI for Difference and P-value				
a. Bootstrap method [2]	[1.8%, 14.6%], 0.0116		[3.4%, 11.3%], 0.0003	
b. GEE with identity link [3]	[1.9%, 14.6%], 0.0113		[3.3%, 11.3%], 0.0003	

[1] All randomized and treated subjects were included in the computation. Missing values were treated as failures.
[2] Based on the Bootstrap method described in Appendix C-1.
[3] Based on a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes were treated as repeated measures.

C-1: Simulation study to assess coverage probabilities of 95% confidence intervals

The reviewer conducted a simulation study to investigate coverage probabilities of the CIs for the difference in the response rates in the Schirmer's test score. Three different 95% CIs were considered:

- The Applicant's 95% CI (based on the GEE with the identity link, but ignored the correlation between the two eyes);
- The 95% CI from the GEE model with the identity link while accounting for the correlation between the two eyes;
- The 95% CI from the Bootstrap method.

In this simulation study, the reviewer used the R package *bindata* (Friedrich Leisch, 2015) to simulate correlated binary responses in the Schirmer's test score from the two eyes of the same subject. Basically, this package creates correlated binary random variables by thresholding normal distributions. The true correlation was assumed to be either 0.4 or 0.5 (the observed correlation was 0.48 in Study 14 and 0.39 in Study 16).

Three different response rates for the OTX-101 0.09% group were assumed: 15%, 16%, or 17% (the observed response rates for the OTX-101 0.09% groups were 16.8% and 16.6% in Study 14 and Study 16, respectively). For each of these three rates, the reviewer considered four different response rates for the vehicle groups: 7%, 8%, 9%, or 10% (the observed response rates for the vehicle groups were 8.6% and 9.2% in Study 14 and Study 16, respectively). For each scenario, 1000 datasets were generated to estimate the coverage probabilities (Each dataset included 744 subjects as Study 16 did).

Table 21 shows the estimated coverage probabilities. As the Applicant's approach ignored the correlation between the two eyes, their 95% CIs had lower coverage than the nominal level of 95%. On the other hand, the coverage for the other two CIs was between 94% and 95%.

Table 21: Coverage probabilities based on 1000 simulated datasets

	Response Rate for OTX-101 0.09%											
	15%				16%				17%			
Correlation =0.4	Response Rate for Vehicle				Response Rate for Vehicle				Response Rate for Vehicle			
	7%	8%	9%	10%	7%	8%	9%	10%	7%	8%	9%	10%
Coverage												
Applicant's CI [1]	0.90	0.90	0.90	0.90	0.90	0.90	0.89	0.90	0.89	0.90	0.90	0.91
GEE CI [2]	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.95	0.95	0.95	0.95
Bootstrap CI [3]	0.94	0.95	0.95	0.95	0.94	0.94	0.94	0.94	0.95	0.95	0.95	0.95
Correlation =0.5	Response Rate for Vehicle				Response Rate for Vehicle				Response Rate for Vehicle			
	7%	8%	9%	10%	7%	8%	9%	10%	7%	8%	9%	10%
Coverage												
Applicant's CI [1]	0.88	0.88	0.88	0.88	0.88	0.88	0.89	0.88	0.88	0.88	0.88	0.89
GEE CI [2]	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.95	0.94	0.94
Bootstrap CI [3]	0.94	0.95	0.95	0.95	0.94	0.95	0.94	0.94	0.95	0.95	0.94	0.95

[1] The Applicant's analysis was equivalent to a simple two-sample Z-test ignoring correlation between left and right eyes;
[2] Based on a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes were treated as repeated measures;
[3] Based on a Bootstrap method with 10,000 Bootstrap samples. Bootstrap samplings were performed at subject-level to account for correlation between the left and right eyes.

Appendix D. Baseline Sign Scores in Study 14

Table 22: Baseline sign scores in OTX-101-2014-001 (study eye and fellow eye)

	Study Eye			Fellow Eye		
	OTX-101 0.05%	**OTX-101 0.09%**	**Vehicle**	**OTX-101 0.05%**	**OTX-101 0.09%**	**Vehicle**
ITT population						
Sample Size	N = 151	N = 152	N = 152	N = 151	N = 152	N = 152
Schirmer's Test Score						
Mean (SD)	10.14 (7.71)	12.24 (9.03)	11.83 (9.00)	11.32 (8.33)	12.44 (9.09)	12.56 (8.93)
Median	9.00	9.00	9.00	9.00	9.00	10.00
Min - Max	0.00 - 35.00	0.00 - 35.00	1.00 - 35.00	0.00 - 35.00	1.00 - 35.00	0.00 - 35.00

(b) (4)

Appendix E. Supportive Efficacy Analyses

Table 23: Proportion of eyes with ≥10 mm increase from baseline in the Schirmer's test score at Day 84 (PP population, missing values were treated as failures)

	OTX-101-2014-001		OTX-101-2016-001	
	OTX-101 0.09%	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Number of eyes	270	268	672	710
Number of eyes with ≥10 mm increase, n (%) [1]	51 (18.9%)	25 (9.3%)	120 (17.9%)	69 (9.7%)
Difference in proportions (OTX-101 - Vehicle)		9.6%		8.1%
95% CI for Difference and P-value				
a. Applicant's analysis [2]	[3.7%, 15.4%], 0.001		[4.6%, 11.8%], < 0.001	
b. Reviewer's analysis [3]	[2.5%, 16.6%], 0.008		[3.9%, 12.5%], < 0.001	

[1] Subjects only in the PP population were included in the computation. Missing values were treated as failures.
[2] The Applicant's analysis was equivalent to a simple two-sample Z-test ignoring correlation between left and right eyes.
[3] Based on a generalized estimating equation model including treatment as a sole factor. Observations from the left and right eyes were treated as repeated measures.

Table 24: Proportion of subjects with ≥ 10 mm increase from baseline in the Schirmer's test score at Day 84 for the average of left and right eyes (ITT population, missing values were treated as failures)

	OTX-101-2014-001		OTX-101-2016-001	
	OTX-101 0.09%	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Number of subjects	152	152	371	373
Number of subjects with ≥10 mm increase, n (%)	25 (16.4%)	11 (7.2%)	48 (12.9%)	26 (7.0%)
Difference in proportions (OTX-101 - Vehicle)		9.2%		6.0%
95% CI for Difference and P-value†	[2.0%, 16.4%], 0.012		[1.7%, 10.3%], 0.006	

† The 95% confidence intervals and p-values were obtained using normal approximations of binomial distributions.

(b) (4)

3 Page(s) has been Withheld in Full as b4 (CCI/TS) immediately following this page

Appendix F. Supportive Summary Tables for AEs

Table 29: AEs leading to study discontinuation

	OTX-101-2014-001			OTX-101-2016-001	
	OTX-101 0.05%	**OTX-101 0.09%**	**Vehicle**	**OTX-101 0.09%**	**Vehicle**
Subjects withdrawn due to AE	4 (4)	5 (3)	6 (1)	17 (12)	3 (2)
Preferred Term					
Instillation site pain	2 (2)	1 (1)	0 (0)	9 (8)	0 (0)
Eye irritation	0 (0)	1 (1)	0 (0)	1 (1)	0 (0)
Eye pain	0 (0)	0 (0)	0 (0)	2 (2)	0 (0)
Back pain	0 (1)	1 (0)	0 (0)	0 (0)	0 (0)
Cerebrovascular accident	0 (0)	0 (1)	1 (0)	0 (0)	0 (0)
Conjunctival hyperaemia	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Conjunctivitis	0 (0)	0 (0)	1 (0)	0 (0)	0 (1)
Conjunctivitis viral	0 (0)	0 (0)	1 (0)	0 (0)	0 (1)
Death	0 (0)	0 (0)	0 (1)	1 (0)	0 (0)
Dyspnoea	0 (0)	1 (0)	0 (0)	0 (1)	0 (0)
Eye allergy	0 (0)	1 (0)	0 (0)	0 (0)	0 (0)
Eyelids pruritus	1 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Foreign body sensation in eyes	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)
Headache	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)
Instillation site discomfort	0 (0)	0 (0)	1 (0)	0 (0)	0 (0)
Nephrolithiasis	0 (0)	0 (0)	0 (0)	1 (0)	0 (0)
Ophthalmic herpes zoster	0 (0)	0 (0)	1 (0)	0 (0)	0 (0)
Scleritis	0 (0)	0 (0)	0 (0)	1 (0)	0 (0)
Seasonal allergy	0 (0)	0 (0)	1 (0)	0 (0)	0 (0)
Throat irritation	0 (0)	0 (0)	0 (0)	1 (0)	0 (0)
Upper respiratory tract infection	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)
Vision blurred	0 (0)	0 (0)	0 (0)	1 (0)	0 (0)

† The number in the parenthesis is the number of AEs suspected to be related to the study medication.

Appendix G. Schedule of Procedures

Table 30: Schedule of procedures in OTX-101-2014-001

Procedures	Screening Days -17 to -15	Run-In Period Days -14 to -1	Baseline Day 0 (± 2)	Day 14 (± 3)	Day 28 (± 3)	Day 42 (± 3)	Day 56 (± 3)	Day 84[a] (+ 7)	Early Discontinuation
Informed consent	X								
Inclusion/exclusion criteria	X		X						
Demographics	X								
Medical/ocular history	X								
Concomitant medication history/review	X		X	X	X	X	X	X	X
Urine pregnancy test[b]□	X							X	X
Run-in period with vehicle		X							
Symptom frequency/severity rating	X		X	X	X	X	X	X	X
Corrected Snellen visual acuity[c]	X		X	X	X	X	X	X	X
Slit lamp examination	X		X	X	X	X	X	X	X
Fluorescein staining of the cornea	X		X	X	X	X	X	X	X
TBUT	X		X	X	X	X	X	X	X
Lissamine green conjunctival staining[d]	X		X	X	X	X	X	X	X
Schirmer's test (unanesthetized)[e]			X					X	X
IOP[f]	X				X		X	X	X
Ophthalmoscopy/dilated fundoscopy	X							X	X
Randomization			X						
Study medication distribution			X		X		X		
Run-in medication distribution	X								
Study medication administration[g]			X					X	
Evaluation of subject comfort and tolerability[h]			X					X	
Patient satisfaction with treatment rating[i]					X		X	X	X
Adverse event assessment[j]			X	X	X	X	X	X	X
Study medication accountability[k]			X		X		X	X	X

[a] Several procedures are performed in a different order on the Day 84 Visit than at the prior visit. Please refer to Protocol Section 10.5 for the order of procedures for this visit.
[b] Women of childbearing potential only
[c] Refraction must be within 6 months of Screening; this refraction should be used for all visual acuity assessments for the duration of the study. The subject must wear the same glasses, if applicable, at each visit.
[d] Lissamine greening conjunctival staining should be performed approximately 5 minutes after corneal staining.
[e] Schirmer's test (unanesthetized) should be performed at least 5 minutes after lissamine green conjunctival staining to allow for any reflex tearing to subside.
[f] Measure IOP using Goldmann applanation tonometry.
[g] The first dose of study medication (Day 0) and the morning dose on Day 84 will be administered by the subject at the site, otherwise subjects will self-administer study medication BID at home. At Day 84, because subjects will dose at the site at this visit, the visit should be scheduled in the morning if possible, and subjects should be reminded at scheduling and confirmation not to dose at home prior to coming to the site. Subjects who are scheduled in the afternoon should be instructed to take their morning dose, but not their second dose prior to coming to the site.
[h] The evaluation of comfort and tolerability will be performed at 3 minutes and 10 minutes after the dose of study medication at the site.
[i] This will be performed prior to visual acuity assessment at the visits indicated.
[j] Collection of AEs extends from signing of informed consent until the last study visit.
[k] Clinical site personnel will document all received and returned medication at Day 0 (run-in medication) and Days 28, 56, and 84 (randomized study medication). Study medication accountability will be performed by the monitor at each applicable monitoring visit.

†Source: Table 2 of the protocol for Study OTX-101-2014-001.

Table 31: Schedule of procedures in OTX-101-2016-001

Procedures Day	Screening -20 to -14	Run-In [a] 14 to 20 days	Baseline 0	Visit 3 28 ± 3	Visit 4 56 ± 3	Visit 5 [b] 84 + 7	Early Discontinuation
Informed consent	X						
Inclusion/exclusion criteria	X		X				
Demographics	X						
Medical/ocular history	X						
Concomitant medication history/review	X		X	X	X	X	X
Urine pregnancy test [c]	X		X			X	X
Run-in period with vehicle		X					
Symptom frequency/severity rating (SANDE)	X		X	X	X	X	X
OSDI questionnaire			X			X	X
Corrected Snellen visual acuity [d]	X		X	X	X	X	X
Slit lamp examination	X		X	X	X	X	X
Fluorescein staining of the cornea	X		X	X	X	X	X
Lissamine green conjunctival staining [e]	X		X	X	X	X	X
Schirmer's test (unanesthetized) [f]			X			X	X
IOP [g]	X					X	X
Ophthalmoscopy/dilated fundoscopy	X					X	X
Randomization			X				
Study medication distribution			X	X	X		
Run-in medication distribution	X						
Study medication administration [h]			X				
Adverse event assessment [i]	X		X	X	X	X	X
Study medication accountability [j]			X	X	X	X	X

[a] The run-in period commences on the day of the Screening Visit. The first dose of vehicle will be instilled at the clinical site.
[b] Several procedures are performed in a different order at Visit 5 than at the prior visits. Please refer to Protocol Section 10.5 for the order of procedures for this visit.
[c] Women of childbearing potential only.
[d] Refraction must be within 6 months of the Screening Visit; this refraction should be used for all visual acuity assessments for the duration of the study. The subject must wear the same glasses, if applicable, at each visit.
[e] Lissamine green conjunctival staining should be performed approximately 5 minutes after corneal staining.
[f] Schirmer's test (unanesthetized) should be performed at least 5 minutes after lissamine green conjunctival staining to allow for any reflex tearing to subside.
[g] Measure IOP using Goldmann applanation tonometry.
[h] The first dose of study medication (Day 0) will be administered by the subject at the site; otherwise subjects will self-administer study medication BID at home.
[i] Collection of AEs extends from signing of informed consent until the last study visit.
[j] Clinical site personnel will document all received and returned medication at Day 0 (run-in medication) and Days 28, 56, and 84 (randomized study medication). Study medication accountability will be performed by the monitor at each applicable monitoring visit.
Abbreviation: SANDE=Symptom Assessment iN Dry Eye.

†Source: Table 2 of the protocol for Study OTX-101-2016-001.

Appendix H. Ocular Surface Disease Index Questionnaire

Ocular Surface Disease Index© (OSDI©)†

Ask your patient the following 12 questions, and circle the number in the box that best represents each answer. Then, fill in boxes A, B, C, D, and E according to the instructions beside each.

HAVE YOU EXPERIENCED ANY OF THE FOLLOWING *DURING THE LAST WEEK*:

	All of the time	Most of the time	Half of the time	Some of the time	None of the time
1. Eyes that are sensitive to light?	4	3	2	1	0
2. Eyes that feel gritty?	4	3	2	1	0
3. Painful or sore eyes?	4	3	2	1	0
4. Blurred vision?	4	3	2	1	0
5. Poor vision?	4	3	2	1	0

HAVE PROBLEMS WITH YOUR EYES LIMITED YOU IN PERFORMING ANY OF THE FOLLOWING *DURING THE LAST WEEK*:

	All of the time	Most of the time	Half of the time	Some of the time	None of the time	
6. Reading?	4	3	2	1	0	N/A
7. Driving at night?	4	3	2	1	0	N/A
8. Working with a computer or bank machine (ATM)?	4	3	2	1	0	N/A
9. Watching TV?	4	3	2	1	0	N/A

HAVE YOUR EYES FELT UNCOMFORTABLE IN ANY OF THE FOLLOWING SITUATIONS *DURING THE LAST WEEK*:

	All of the time	Most of the time	Half of the time	Some of the time	None of the time	
10. Windy conditions?	4	3	2	1	0	N/A
11. Places or areas with low humidity (very dry)?	4	3	2	1	0	N/A
12. Areas that are air conditioned?	4	3	2	1	0	N/A

†Source: Appendix 1 of the protocol for Study OTX-101-2016-001.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

WONYUL N LEE
07/05/2018

YAN WANG
07/05/2018
I concur.