Luminopia

LUMINOPIA ONE Amblyopia Vision Improvement Study

Statistical Analysis Plan

Table of Contents

1 Introduction

This statistical analysis plan (SAP) describes the planned statistical methods to be used during the reporting and analysis of data collected under the Clinical Investigation Protocol Luminopia One Amblyopia Vision Improvement Study. This SAP should be read in conjunction with case report forms (CRFs). This version of the SAP has been developed with respect to the Clinical Investigation Protocol Luminopia One Amblyopia Vision Improvement Study Version 2.0 Amendment 1 08Oct2018. Any revisions to the protocol or CRFs that impact the planned analyses may require updates to the SAP. Analyses not specified in this SAP will be considered post-hoc and exploratory in nature. The SAP considers VA values in true logMAR scores; the protocol reports VA in lines, where 1 line is equal to 0.10 logMAR.

2 Abbreviations

Abbreviation/Term	Definition
AE	Adverse Event
AT	As-Treated
BCVA	Best Corrected Visual Acuity
CIP	Clinical Investigation Plan
CRF	Case Report Form
eCRF	Electronic Case Report Form
ICF	Informed Consent Form
ITT	Intent-to-Treat
PP	Per-Protocol
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
VA	Visual Acuity

3 Study Objectives

The objective of the study is to demonstrate the safety and efficacy of Luminopia One in patients with amblyopia associated with anisometropia and/or with mild strabismus.

3.1 Primary Efficacy Endpoint

The mean change in amblyopic eye BCVA from baseline to 12 weeks of treatment with Luminopia One compared to mean change in amblyopic eye BCVA from baseline to 12 weeks with refractive correction alone.

3.2 Primary Safety Endpoint

The mean change in fellow eye BCVA from baseline to 12 weeks of treatment with Luminopia One compared to mean change in fellow eye BCVA from baseline to 12 weeks with refractive correction alone.
Additionally, the primary safety endpoint will report the frequency and severity of all related Adverse Events (anticipated and unanticipated).

3.3 Secondary Endpoints

3.3.1 VA at 12 Weeks Defined as a Binary Outcome

The proportion of amblyopic-eye BCVA improvement of ≥ 2 lines (-0.20 logMAR) at 12 weeks (Visit 4) from baseline, in each treatment group.

3.3.2 Stereoacuity at 4, 8 and 12 Weeks

Mean change in stereoacuity from baseline to 4, 8 and 12 weeks, respectively, in each treatment group.

3.3.3 VA at 4 and 8 Weeks

The mean change in amblyopic eye BCVA from baseline to 4 and 8 weeks of treatment with Luminopia One compared to mean change in amblyopic eye BCVA from baseline to 4 and 8 weeks with refractive correction alone.

3.3.4 Adherence at 4, 8 and 12 Weeks

The mean adherence with the Luminopia One therapeutic from baseline to 4, 8, and 12 weeks, in the treatment group.

3.4 Safety Reporting

3.4.1 Adverse Events

The participant and parent(s) will be asked about anticipated and unanticipated adverse events at each visit and phone call. The frequency of known and unknown adverse events will be summarized for all participants.

4 Study Design

The proposed study is a multi-center randomized, controlled, masked, clinical trial which compares mean change in amblyopic eye BCVA from baseline with Luminopia One (“therapeutic”), to refractive correction (“control”). One-hundred and forty participants (n = 140) aged 4-7 years (inclusive) will be enrolled. Participants will be randomized 1:1 to the “therapeutic group”, to use Luminopia One, or the “control group”, to undergo continued refractive correction, for 12 weeks of treatment.

A group sequential design with a single interim analysis will be used to allow stopping for either early success or futility in both primary safety and efficacy endpoints. The interim analysis will be conducted after 75% of participants have completed their 12-week follow-up visit. The overall Type I error rate was set at 0.05.

4.1 Randomization

Participants will be randomized at the screening and enrollment visit. Randomization will be performed using a random permuted block design stratified by investigational site with a 1:1 allocation ratio (therapeutic group : control group). Randomized treatment will be assigned from the electronic database randomization assignment module.

4.2 Masking

Due to the nature of the therapy, participants cannot not be masked to the treatment assignment. Stereoacuity and distance visual acuity testing will be conducted by a Masked Examiner who will not have access to treatment group assignment.

5 Sample Size Determination

Sample size estimates were based on data from previous Pediatric Eye Disease Investigator Group (PEDIG) studies (ATS3[1], ATS5[2], ATS18[3] and ATS20) and data from participants in Luminopia One pilot trials who would meet the eligibility criteria for the current protocol.

Control Group –Refractive Correction

To estimate the mean change in amblyopic eye visual acuity for the primary efficacy endpoint, data were reviewed from participants aged 4 to 6 years who were randomly assigned to continue refractive correction alone in a previous PEDIG study (ATS5, see Table 1). The mean change in VA at 5 weeks was -0.055 logMAR (95% CI: -0.019 to -0.090 logMAR) with standard deviation 0.128 logMAR (95% CI: 0.107 to 0.158 logMAR). A negative change in logMAR score of VA from baseline to 5 weeks indicates an improvement in vision.

Given the more stringent eligibility criteria in the proposed study as compared to the previous study, we anticipate that the magnitude and standard deviation of mean change in visual acuity after 12 weeks with refractive correction in the current study will be smaller than those in Table 1. The anticipated improvement in the control group, for the purpose of sample size estimation, was set at -0.045 logMAR by a group of PEDIG investigators and the Sponsor.

To estimate the mean change in fellow eye visual acuity for the primary safety endpoint, data were reviewed from participants who were randomly assigned to continue refractive correction alone in a previous PEDIG study (ATS5, see Table 2). The mean change in VA at 5 weeks was -0.02 logMAR (95% CI: 0.00 to -0.04 logMAR) with standard deviation 0.10 logMAR (95% CI: 0.08 to 0.12 logMAR). A negative change in logMAR score of VA from baseline to 5 weeks indicates an improvement in vision.

Therapeutic Group – Luminopia One

To estimate the mean change in amblyopic eye visual acuity for the primary efficacy endpoint, data from 3 studies were reviewed: (1) data for participants randomly assigned to a similar treatment of “Dig Rush” on an iPad device (E. Birch Study); (2) data from a previous PEDIG trial (ATS18) for participants randomly assigned to a similar treatment of “Hess falling blocks” on an iPad device; and (3) data from Luminopia One pilot trials (Table 1).

To estimate the mean change in fellow eye visual acuity for the primary safety endpoint, data were reviewed from participants aged 4 to 7 years enrolled in Luminopia One pilot trials. The mean change in VA at 12 weeks was -0.05 logMAR (95% CI: -0.02 to -0.08 logMAR) with standard deviation 0.10 logMAR

(95% CI: 0.08 to 0.13 logMAR). A negative change in logMAR score of VA from baseline to 12 weeks indicates an improvement in vision.

Table 1: Previous Study Data (Primary Efficacy Endpoint Reference)

Cohort	N	Change in Amblyopic Eye VA at the 4 -5 Week Visit (dependent on study design)		N	Change in Amblyopic Eye VA at 8-12 Week Visit (dependent on study design) ‡	
		Mean logMAR) Change (95% CI)	SD Change (95% CI)		Mean logMAR)Change (95% CI)	SD Change (95% CI)
Refractive Correction Alone:						
ATS5: Age 3 to 6 years [1]	53	-0.055 (-0.020 to -0.089)	0.128 (0.107 to 0.158)		n/a	n/a
Binocular Therapeutic:						
E. Birch Study (Age 4 to 6 years) [2]	9	-0.167 (-0.089 to -0.245)	0.119 (0.080 to 0.228)		n/a	n/a
Luminopia Pilot Trials (Age 4 to 7 years) [3]	34	-0.136 (-0.096 to -0.176)	0.118 (0.095 to 0.155)	35	-0.164 (-0.109 to -0.219)	0.166 (0.134 to 0.217)
ATS18 (Age 5 to 6 years) [4]	39	-0.115 (-0.073 to -0.157)	0.135 (0.110 to 0.174)	39	-0.123 (-0.079 to -0.167)	0.139 (0.114 to 0.179)
Patching Therapeutic:						
ATS18 (Age 5 to 6 years) [4]	50	-0.090 (-0.050 to -0.130)	0.145 (0.121 to 0.181)	48	-0.143 (-0.098 to -0.188)	0.162 (0.135 to 0.202)

[1] ATS5 data were limited to randomized participants with an amblyopic-eye VA of 20/40 to 20/200 inclusive with ≥ 3 lines of interocular difference and a fellow-eye VA of 20/25 or better who had stabilized with refractive correction prior to randomization. The baseline magnitude of tropia at near (as measured by SPCT) was limited to <5pd. The outcome data reported were based on the 5-week primary masked outcome visit.

[2] E. Birch study enrolled children aged 4 to <7 years of age. Change in VA after 4 weeks of binocular therapy was computed for participants randomly assigned to receive binocular the Dig Rush game on an iPad device for 4 weeks (prescribed 1 hour per day, 5 days per week without patching). Participants with >4pd magnitude of strabismus (measured by PACT) were excluded from the study

[3] Data from Luminopia pilot trials are limited to participants 4 to 7 years old with ocular alignment ≤5pd.

[4] ATS18 study participants were prescribed binocular therapy for 1 hour per day, 7 days per week vs 2-hours per day of patching. The binocular game, Hess falling blocks, was played on an iPad device. The outcome data reported are from the 4-week and 8-week masked outcome visits.

Table 2: Previous Study Data (Primary Safety Endpoint Reference)

Cohort	N	Change in Fellow Eye VA at the 4-5 Week Visit (logMAR)		N	Change in Fellow Eye VA at 12 Week Visit (logMAR) ‡	
		Mean Change (95% CI)	SD Change (95% CI)		Mean Change (95% CI)	SD Change (95% CI)
Refractive Correction Alone:						
ATS5: Age 3 to 6 years [1]	68	-0.02 (-0.00 to -0.04)	0.10 (0.08 to 0.12)		n/a	n/a
Binocular :						
Luminopia Pilot Trials (Age 4 to 7 years) [3]	36	0.00 (-0.02 to 0.02)	0.05 (0.04 to 0.06)	35	-0.05 (-0.02 to -0.08)	0.10 (0.08 to 0.13)

Effect Size Estimates

Efficacy:

Based on Luminopia One pilot trials (Table 1), the estimate for mean change in amblyopic eye visual acuity for the primary efficacy endpoint is -0.164 logMAR in the therapeutic group and -0.045 logMAR in the control group, with a pooled standard deviation of 0.140 logMAR. The pooled standard deviation was estimated from two studies run by PEDIG in the same patient population ATS 18[4] and ATS 20[5].

The true difference between the therapeutic and refractive correction might be as large as -0.119 logMAR; however, a conservative approach was taken to detect a true difference as small as -0.075 logMAR, and therefore a group difference of -0.075 logMAR will be used to estimate sample size. There is sufficient evidence supporting the therapeutic's efficacy to justify a one-sided superiority test.

Safety:

Safety will be measured as change in visual acuity in the fellow eye. This is a noninferiority analysis to demonstrate that change of VA in the fellow eye when using Luminopia One is not worse than that observed in the control group.

Based on previous studies of amblyopia where Fellow Eye VA was the safety endpoint (Table 2), the estimate for mean change in fellow eye visual acuity for the primary safety endpoint is -0.05 logMAR in the therapeutic group and -0.02 logMAR in the control group, with a pooled standard deviation of 0.10 logMAR.

A noninferiority margin of +0.10 logMAR (1 line of VA) was determined to be most appropriate, given that 0.10 logMAR is the minimum clinically detectable difference in clinical practice with a standard eye chart for this age group[38]. Consequently, +0.10 logMAR would be the smallest unacceptable difference in fellow eye visual acuity change between the therapeutic and control groups, and therefore the most appropriate noninferiority margin.

Additionally, the primary safety endpoint will report the frequency and severity of all related Adverse Events (anticipated and unanticipated).

Group Sequential Design:

The sample size for this trial will be based on a group sequential design with a single interim analysis to allow for stopping due to either early success (rejecting the null hypothesis) or futility. The goal of the group sequential design is to conserve precious time and resources and prevent unnecessary enrollment of participants if there is no observed difference in effect between groups **or** if there is a significant effect in favor of Luminopia One. It is important to allow for early stopping for success since there is a real possibility the therapeutic could perform significantly better than expected.

Boundaries for both alpha and beta will be calculated using a power cumulative error spending function. The mathematical form of this spending function is:

$$E(t,\rho) = \begin{cases} 1 \, if \; t \geq 1 \\ t^{\rho} \; if \; 0 < t < 1 \\ 0 \, otherwise \end{cases}$$

where ρ is the power parameter and t is the information fraction.

For beta spending, rho was set to 1.0 which results in a spending function much like the Pocock method (i.e. equal spending between the interim and final analyses). For alpha spending, rho was set to 2.0 resulting a spending function somewhere between that found for the O'Brien-Fleming method and the Pocock method.

The planned interim analysis will also involve an analysis of the primary safety endpoint, including both the test of noninferiority in mean fellow eye BCVA change between groups, and a report of the frequency and severity of all related Adverse Events (anticipated and unanticipated).

<u>Sample Size Estimation</u>

The sample size was calculated using the SEQDESIGN procedure in SAS version 9.3. The alternative reference was set to 0.075 logMAR (0.75 lines). As part of the design statement, the number of analyses was set to 2 where the first analysis was planned after 75% of the expected participants complete their 12-week primary outcome visit, and the second after all participants complete their 12-week primary outcome visit. The method of alpha spending was set to "errfuncpow(rho=2.0)" and the method of beta spending was set to "errfuncpow(rho=1.0)." The "alt" parameter was set to "upper" to indicate a one-sided test, alpha was set to 0.05, and beta was set to 0.10 in order to achieve 90% power. The "stop" parameter was set to "both" in order to have boundaries for both success and futility.

This specification results in an effective sample size of 132 participants (66 per arm) providing 90% power to detect a 0.075 logMAR difference between the therapeutic and control groups using a one-sided test (primary **efficacy** endpoint).

With an equal beta-spending function, if there truly is no treatment effect favoring Luminopia One, the chance of stopping the study early for futility at the interim analysis is approximately 89%.

The primary **safety** endpoint will test noninferiority in fellow eye visual acuity change from baseline to the 12-week primary endpoint. The sample size required to achieve 90% power, for a non-inferiority window of 1.0 line of VA (-0.10 logMAR) with Type I error rate of 0.05, is fewer than 132 participants. Therefore, the primary efficacy endpoint will be used to drive sample size.

Adjusting the effective sample size for a 5% loss to follow-up results in a target enrollment of **140 participants**.

6 Statistical Analyses

6.1 General Considerations

Unless otherwise specified, the following general principles apply to the planned statistical analyses. All statistical analysis will be conducted using SAS version 9.4 or later (SAS Institute Inc., Cary, NC) or other widely-accepted statistical or graphical software as required.

6.1.1 Descriptive Statistics

Continuous data will be summarized with mean, standard deviation, median, minimum, maximum, and number of evaluable observations. Categorical variables will be summarized with frequency counts and percentages. Confidence intervals may be presented, where appropriate.

6.1.2 Study Day

Study day 0 is the date of the randomization. For each participant, duration in study will be based on date of randomization and the last study contact date which is the latest date of all follow-up visits, assessments, adverse event onset or resolution, and study exit including date of death.

6.1.3 Visit Windows

The primary endpoint analyses on mean change in BCVA (in both amblyopic and fellow eyes) will be based on in-window visits only. An additional sensitivity analysis will be conducted using all available data associated with the 12-week visit regardless of in or out of window. Unless otherwise specified, all secondary and exploratory visit-based assessments will be analyzed for each analysis time point according to the nominal visit entered in the Case Report Form (CRF) regardless of if it is out of window. See section 3.4 in the protocol for visit window specifications.

6.1.4 Statistical Significance

The primary efficacy and safety endpoints will performed using one-sided tests and a 0.05 significance level. Any additional exploratory hypothesis testing will be performed at the two-sided 0.05 significance level. P-values will be rounded to three decimal places. If a p-value is less than 0.001 it will be reported as “<0.001”. If a p-value is greater than 0.999, it will be reported as “>0.999”.

6.2 Analysis Populations

The following analysis populations are defined for analysis:

1. **Intention-to-treat (ITT) Analysis Set:** The intention-to-treat (ITT) population consists of all enrolled participants who were subsequently randomized. The ITT set will be analyzed according to randomized treatment assignment regardless of treatment actually received. The ITT population will not include participants who did not meet eligibility criteria at the time of enrollment and are ultimately deemed ineligible for the study, including participants who are withdrawn from the study due to insufficient internet.
2. **As-Treated Analysis Set:** The as-treated (AT) population consists of participants enrolled and randomized, however treatment group assignment for the purposes of analysis will be based on the actual treatment received. Participants randomized to the Luminopia One group with ≤ 25% adherence in device use at 12 weeks will be analyzed as control participants.
3. **Per-Protocol Analysis Set:** The per-protocol (PP) population consists of all ITT participants who were treated according to randomized assignment, who had adherence > 25% in Luminopia One Use for the therapeutic group, and who had no major protocol deviations that could affect the endpoints of interest. Protocol deviations will be reviewed and defined as major prior to analysis

of study outcomes. The 25% adherence threshold for the PP population was determined from pilot trials on Luminopia One, in which all enrolled participants with > 25% adherence and no major protocol deviations showed improvements from using the therapeutic. Adherence will be captured while maintaining masking of the study in accordance with the Masking Plan.

The ITT population will be used for analysis and interpretation of the primary and secondary efficacy endpoints and the co-primary safety endpoint assessing BCVA in the fellow eye. Supportive analysis for the primary and secondary efficacy endpoints will be performed using the PP population to provide insight into the potential impact of protocol deviations on the primary results of the study. The AT cohort will be used to characterize and assess safety, including the co-primary safety endpoint evaluating frequency and severity of all related Adverse Events, as well as supportive efficacy analyses.

6.3 Handling of Missing Data

Multiple imputation for missing data will be performed as a secondary approach with the ITT population, and results of the analysis with imputation of missing data will be considered supportive in nature and assessed for consistency with the primary analysis. Imputation will be based on regression methods using baseline covariates as independent predictors. If no significant predictors can be found, the following method will be used. Those participants completing the study will be categorized by age, sex, and treatment received. For a participant with missing data, the imputed value will be a random draw from those participants with data that have the same age, sex, and treatment. See Section 6.6 for imputation methods specific to the study endpoints.

In the case of partial dates (e.g. adverse event onset date), the unknown portion of the date of the event will be imputed. If the month and year are known, the 15th of the month will be used for analysis. If only the year is known, the event will be analyzed as if it occurred on June 30th of the known year. Imputation of partial dates is participant to the condition that it must occur on or after the randomization date. In the case where the imputed date is prior to the randomization date, the date of the randomization will be used.

6.4 Participant Disposition

The number of participants in each analysis population will be presented along with reason for any exclusions. Participant accountability will be summarized by visit. The number of participants who are enrolled, eligible for follow-up, and number completing clinical follow-up will be summarized for each protocol-required visit. In addition, the number of participants who complete the study or exit early will be summarized by reason.

6.5 Demographics and Baseline Characteristics

Descriptive statistics will be presented for all clinically-relevant baseline demographic and medical history variables.

6.6 Analysis of Study Endpoints

6.6.1 Primary Efficacy Endpoint

The mean change in amblyopic eye BCVA from baseline to 12 weeks of treatment with Luminopia One compared to mean change in amblyopic eye BCVA from baseline to 12 weeks with refractive correction alone.

6.6.1.1 Primary Analysis

An analysis of variance will be performed to compare the change in VA from baseline to 12 weeks in the therapeutic and control groups. The primary objective will be assessed with the following hypothesis:

H_0: $\beta_1 \geq 0$

H_a: $\beta_1 < 0$

where β_1 is the coefficient from the following model:

$Y = \beta_0 + \beta_1 * X_1$,

where Y is the adjusted average change in BCVA in the amblyopic eye from baseline to 12 weeks, β_1 and X_1 are the coefficient and regression term for treatment arm (X_1 coded as therapy=1, control=0), respectively.

The hypothesis will be evaluated using a one-sided F-test for the coefficient associated with treatment group. A standard F-test will be used to determine if there is a significant difference in mean change from baseline between groups. The endpoint will be presented as average BCVA at baseline and 12 weeks, change in BCVA from baseline to 12 weeks in each treatment group, difference in BCVA change between treatment groups, two-sided 90% confidence interval of the difference, and p-value associated with the hypothesis test above. The endpoint will be considered met if the p-value for the coefficient is <0.05.

The endpoint will be evaluated using the ITT population with no imputation for missing data. Assessments from out of window visits will be considered missing in the primary analysis. Supplementary analysis will be provided based on ITT with imputation, PP and AT cohorts.

The 24-week visit for participants in the control group who switch over to the therapeutic at the 12-week visit will not be a part of the primary analysis.

6.6.1.2 Handling of Missing Data

Missing data will not be imputed in the primary analysis; participants missing either baseline BCVA or 12-week BCVA, or with an out-of-window 12-week visit will not be included. A sensitivity analysis will be performed to assess the impact of missing data. A supportive analysis will be conducted including available data from out of window visit assessments.

6.6.1.3 Sensitivity Analysis

A sensitivity analysis will be conducted using multiple imputation for missing data and results of the analysis will be assessed for consistency with the primary analysis. Imputation will be based on regression methods using significant baseline covariates as independent predictors. Covariates will be determined based on a statistically or clinically significant association with the change in BCVA.

Missing data will be imputed using the PROC MI procedure in SAS. PROC MI imputes missing covariates in the order variables are listed on the VAR statement. The variables will be listed on the VAR statement in the order listed above.

Missing covariates will be imputed using Markov Chain Monte Carlo (MCMC) methods. For the endpoint, the method will be determined based on the pattern and nature of the missing data (e.g. FCS regression method for an arbitrary missing pattern). The imputation will be executed in two steps in SAS. The first step will use PROC MI to impute the missing covariates. For each imputed dataset generated in step 1, the endpoint status will be imputed in step 2 to create the final imputed datasets.

Explorations to omit predictors may be conducted if the multiple imputation models will not converge. For each multiple imputation model, 10 imputed datasets will be generated. The imputed datasets will then be combined for inference using standard methods such as those available in SAS PROC MIANALYZE or other valid statistical software.

If no significant predictors are identified, the following method for imputation will be used. Participants who completed the study will be categorized by age, sex, and treatment received. For a participant with missing data, the imputed value will be a random draw from those participants with data that have the same age, sex, and treatment.

Additional sensitivity analyses may be performed if deemed necessary.

6.6.2 Primary Safety Endpoint

The co-primary analysis of safety will test if the mean change in fellow eye visual acuity of the therapeutic group from baseline to 12 weeks is non-inferior to the mean change in fellow eye visual acuity of control group within a margin of 0.10 logMAR. The formal hypothesis being tested are:

$$H_0: \mu_{therapeutic} - \mu_{control} \geq \delta$$

H_a: $\mu_{therapeutic} - \mu_{control} < \delta$

where μ is the change from baseline (follow-up - baseline) for both the control and therapeutic participants and δ is the non-inferiority window of 0.10 logMAR. The statistical test is a two-sample one-sided T-test which will be done at the 0.05 level of significance. An upper 95% confidence limit for the difference will also be presented. The endpoint will be considered met if the p-value is < 0.05.

The endpoint will be evaluated using the ITT population with no imputation for missing data. Assessments from out of window visits will be considered missing in the primary analysis. Supplementary analysis will be provided based on ITT with imputation, PP and AT cohorts, as described in the primary efficacy endpoint analysis.

In addition to the formal test on the mean change in fellow eye visual acuity, the primary safety endpoint will report the frequency and severity of all related Adverse Events (anticipated and unanticipated). Count of events, percent of participants who experienced an event, and an exact 95% confidence interval of the event rates will be presented for each treatment group by event type. No formal hypothesis testing will be performed. Event rates will be summarized for clinical interpretation. Safely analyses on Adverse Events will be based on the AT analysis cohort.

6.6.3 Study Success

The study will be considered a success if both the primary efficacy and primary safety endpoints are passed. The 12-week visit measurements in the therapeutic and control groups will determine success and failure of the trial for both primary safety and efficacy endpoints. Measurements from the 24-week visit, for participants in the control group who switch over to the therapeutic after the 12-week visit will not be part of the primary analysis and will not define study success.

6.6.4 Secondary Efficacy Endpoints

No formal hypothesis testing will be performed for the secondary endpoints; any statistical comparisons made will be considered exploratory and no adjustments for multiple testing will be made.

The secondary endpoints will be evaluated using the ITT population with no imputation for missing data. Supplementary analysis may be provided based on the PP and AT cohorts. Sensitivity analyses, including missing data imputation, may be performed if necessary. Unless otherwise stated, only in-window follow-up visits will be included in the endpoint analyses.

6.6.4.1 VA at 12 Weeks of 2.0 or more lines

The proportion of participants with amblyopic-eye BCVA improvement of ≥ 2 lines of VA (≤ -0.2 logMAR) at 12 weeks (Visit 4) from baseline will be presented by group and an exact 95% confidence interval will

also be provided. Subjects with available data at baseline and an in-window 12-week visit will be included in the analysis.

6.6.4.2 Stereoacuity at 4, 8, and 12 Weeks

A secondary analysis will report the mean change in stereoacuity in each treatment group from baseline to 4, 8 and 12 weeks along with 95% confidence intervals about the mean change. Subjects with available data at baseline and an in-window follow-up visit will be included in the analysis.

Stereoacuity is evaluated with the Titmus Fly and Randot Preschool tests. The final stereoacuity measurement used for analysis for each visit is the lower, i.e. better, score between the two tests. The log transformation of the final score is used for analysis and back-transformed for displaying summary results.

6.6.4.3 VA at 4 and 8 Weeks

The primary endpoint analysis will be repeated for mean change in amblyopic eye BCVA from baseline to 4 and 8 weeks of treatment with Luminopia One will be compared to continued refractive correction. Mean change and 95% confidence intervals will be presented. Sensitivity and supplementary analyses may be performed as needed. Subjects with available data at baseline and an in-window follow-up visit will be included in the analysis.

6.6.4.4 Adherence at 4, 8, and 12 Weeks in the Luminopia One Group

A secondary analysis will report the mean adherence from baseline to 4, 8, and 12 weeks along with 95% confidence intervals about the mean adherence in refractive correction wear, and at 12 weeks in Luminopia One use.

6.6.5 Secondary Safety Endpoints

Safety analyses will be performed on the AT population. Results from any hypothesis testing will be considered exploratory and no adjustments for multiple testing will be made.

6.6.5.1 Adverse Events

AEs will be tabulated with the number of events and participants with event for each event type as described in the primary safety analysis. Additional summaries may be provided as necessary. Rates will be reported as the number of participants who experience at least one event during the analysis interval out of the total number of participants.

6.7 Subpopulation and Exploratory Analyses

The mean change in amblyopic eye BCVA from baseline to 12 weeks will be evaluated in subpopulations based on baseline factors in exploratory analyses and used to suggest hypotheses for further

investigation in future studies. The following baseline factors are of interest: prior amblyopia treatment vs no prior amblyopia treatment, type of amblyopia (amblyopia without vs with strabismus), age (≤5 vs > 5), and baseline VA where the range of baseline VA is split into two groups (better than 20/100 vs 20/100 or worse). The general approach for these exploratory analyses will be to conduct an analysis similar to the primary analysis adding an interaction for treatment and the subgroup covariate of interest. If the p-value for the interaction term is < 0.15, additional exploratory analysis may be performed to understand any variations in the outcomes by subgroup.

Additional subgroups may be defined during the analysis phase for investigative purposes. Exploratory analyses will also evaluate the relationship between adherence and mean change in amblyopic eye VA at the primary endpoint visit in the therapeutic group and will evaluate the change in amblyopic-eye VA from the 12-week visit to the 24-week visit for participants who were initially randomized to the control group and switched over to the therapeutic at the 12-week visit.

6.8 Site Poolability Analyses

Data from this trial will be pooled across study site for the purposes of analysis. To demonstrate poolability of the primary endpoint data, the change in amblyopic eye BCVA between baseline and 12 weeks will be calculated for each participant. This value will be used as the dependent variable in a linear model. The independent predictors will be study site, treatment group, and a study site by treatment group interaction. If the interaction term is considered non-significant then the data will be pooled across site for the purposes of analysis. If the interaction term is considered significant then the analysis of the primary endpoint will have to be adjusted for study site. A significance level of 0.15 will be used for this analysis. Sites with fewer than 5 participants enrolled may be combined into one site for purposes of this analysis.

6.9 Interim Analyses

6.9.2 Group Sequential Design

A group sequential design will be employed to allow for early stopping due to either success (rejecting the null hypothesis) or futility. The goal of the group sequential design is to conserve precious time and resources and prevent unnecessary enrollment of participants if there is no observed difference in effect between groups or if there is a significant effect in favor of Luminopia One. It is important to allow for early stopping for success since there is a real possibility the therapeutic could perform significantly better than expected.

Boundaries for both alpha and beta will be calculated using a power cumulative error spending function. The mathematical form of this spending function is:

$$E(t,\rho) = \begin{cases} 1 \, if \, t \geq 1 \\ t^{\rho} \, if \, 0 < t < 1 \\ 0 \, otherwise \end{cases}$$

where ρ is the power parameter and t is the information fraction.

For beta spending, rho was set to 1.0 which results in a spending function much like the Pocock method (i.e. equal spending between the interim and final analyses). For alpha spending, rho was set to 2.0 resulting a spending function somewhere between that found for the O'Brien-Fleming method and the Pocock method.

The planned interim analysis will be conducted after 75% of participants, of the re-estimate sample size, reach their 12-week visit and will also involve an analysis of the primary safety and efficacy endpoint, and a report of the frequency and severity of all related Adverse Events (anticipated and unanticipated). All endpoint analyses will be conducted simultaneously on the same data snapshot.

<u>Minimizing Bias</u>

The interim analysis will be designed to optimize usage of time and resources while controlling the Type I and II error rates and minimizing operational bias, in accordance with FDA's guidance on Adaptive Designs for Medical Devices.

Type I and II Error Rates

Analyzing interim data during the study can lead to Type I error inflation and increase the chance of a false positive conclusion. The proposed group sequential design controls for Type I error inflation by using the well-established statistical method of alpha-spending, so the Type I error for the overall study is controlled at an acceptable level. A similar approach of a beta-spending function is used to keep the Type II error for the overall study at an acceptable level.

Operational Bias

The details of the group sequential design will be approved prior to collection of study data. To avoid any change in behavior between study personnel and participants before and after the interim analysis, the interim analysis will be conducted by the CRO and Sponsor without the involvement of participating investigators. The Medical Monitor will be informed of the interim analysis findings.

In the event that the study is stopped early for success or futility, randomization will be halted and the Medical Monitor will be asked to make a recommendation as to whether follow-up exams for existing participants should continue.

6.10 Protocol Deviations

Protocol deviations will be summarized for all deviations and by type with counts and number of participants with at least one deviation.

7 Changes from Planned Analyses

Any changes to planned statistical analyses determined necessary prior to performing the analyses will be documented in an amended Statistical Analysis Plan and approved prior to the analysis when possible. Any other deviations or changes from the planned analyses deemed necessary due to violation of critical underlying statistical assumptions, data characteristics, or missing data will be clearly described in the clinical study report with justification and rationale.

8 References

1. Pediatric Eye Disease Investigator Group. Randomized trial of treatment of amblyopia in children aged 7 to 17 years. *Arch Ophthalmol* 2005;123:437-47.

2. Pediatric Eye Disease Investigator Group. A Randomized Trial to Evaluate Two Hours of Daily Patching for Amblyopia in Children. *Ophthalmology*. 2006;113(6):904-912. doi:10.1016/j.ophtha.2006.01.069.

3. Holmes JM, Manh VM, Lazar EL, Beck RW, Birch EE, Kraker RT, Crouch ER, Erzurum SA, Khuddus N, Summers AI, Wallace DK, for the Pediatric Eye Disease Investigator Group. Effect of a Binocular iPad

Game vs Part-time Patching in Children Aged 5 to 12 Years With Amblyopia. *JAMA Ophthalmol* 2016;134(12):1391-1400.

4. Study of Binocular Computer Activities for Treatment of Amblyopia – Public Dataset https://public.jaeb.org/pedig/stdy/235

5. Binocular Dig Rush Game Treatment for Amblyopia – Public Dataset https://public.jaeb.org/pedig/stdy/506