

**Statistical Analysis Plan**

Sponsor:	GenSight Biologics
Protocol No:	GS-LHON-CLIN-03B (REVERSE)
Version No./ Date:	V4.0/07 <sup>th</sup> November 2016
Title	A Randomized, Double-Masked, Sham-Controlled, Pivotal Clinical Trial to Evaluate the Efficacy of a Single Intravitreal Injection of GS010 (rAAV2/2- ND4) in Subjects Affected for More than 6 Months and To 12 Months by Leber Hereditary Optic Neuropathy Due to the G11778A Mutation in the Mitochondrial NADH Dehydrogenase 4 Gene
CRF Version No./ Date	V6.0/05 July 2017
PRA Project Id:	GNSGS005-GS005X
SAP Version No./ Date:	SAP Amendment Version 1.0/12 <sup>th</sup> March 2018

<b>Approvals .....</b>	<b>1</b>
<b>2. Study Objectives.....</b>	<b>4</b>
<b>3. Study Design .....</b>	<b>5</b>
3.1 Overall Design.....	5
3.2 Duration and Schedule of Study Participation.....	6
3.3 Sample Size Considerations.....	7
3.4 Randomization.....	8
<b>4. Study Variables and Covariates .....</b>	<b>9</b>
4.1 Primary Variable.....	9
4.2 Secondary Variables .....	9
4.2.1 Efficacy.....	9
4.2.2 Safety .....	10
4.3 Predetermined Covariates and Prognostic Factors .....	11
<b>5. Definitions .....</b>	<b>11</b>
5.1 Study Day.....	11
5.2 Baseline .....	11
5.3 Change from Baseline .....	11
5.4 Derived Variables.....	11
5.5 Handling of missing data and/or invalid data and outliers .....	14
5.5.1 Efficacy.....	14
5.5.2 Safety .....	15
<b>6. Analysis Sets.....</b>	<b>16</b>
6.1 Intention-to-Treat.....	16
6.2 Per-Protocol population .....	17
6.3 Safety.....	17
<b>7. Interim Analyses .....</b>	<b>17</b>
7.1 Timing of Primary Analysis .....	17
<b>8. Data Review .....</b>	<b>17</b>
8.1 Data Handling and Transfer.....	17
8.2 Data Screening.....	17
<b>9. Statistical Methods.....</b>	<b>18</b>

10. Validation ..... 31

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] P

## 2. STUDY OBJECTIVES

### Primary Objective:

- To evaluate the efficacy of GS010 compared with sham at Week 48 in the change from baseline of the Log of the Minimal Angle of Resolution (LogMAR) in subjects affected for more than 6 months and up to 12 months by LHON.

### Secondary Objectives:

- To assess the effect of GS010 on parameters measured with high resolution spectral-domain optical coherence tomography (SD-OCT), including Retinal Ganglion Cell (RGC) layer, inner and outer nuclear layer (INL and ONL) thickness/volume, Retinal Nerve Fiber Layer (RNFL) thickness and topographical map analysis of these anatomical regions.

- To evaluate the efficacy of GS010 compared with sham on the visual acuity (LogMAR) in a dynamic perspective through the comparison of the time course of the response.
- To verify whether the efficacy at Week 48 and at Week 96 of GS010 compared with sham and measured by the change from baseline of the LogMAR is dependent upon the treatment of the best- or worst-seeing-eye.
- To verify whether the rate of responders at Week 48 and 96 is dependent upon the treatment received and whether the magnitude of the treatment effect is dependent upon the treatment of the best- or worst-seeing eye at entry.
- To assess the effect of GS010 on standardized automated visual fields obtained with the Humphrey Visual Field (HVF) Analyzer II.
- To assess the effect of GS010 on contrast sensitivity measured with the Pelli-Robson chart.
- To assess the effect of GS010 on color vision measured with the Farnsworth-Munsell 100 Hue color test.

[REDACTED]

■

[REDACTED]

[REDACTED]

■

[REDACTED]

■

[REDACTED]

■

[REDACTED]

### 3. STUDY DESIGN

#### 3.1 OVERALL DESIGN

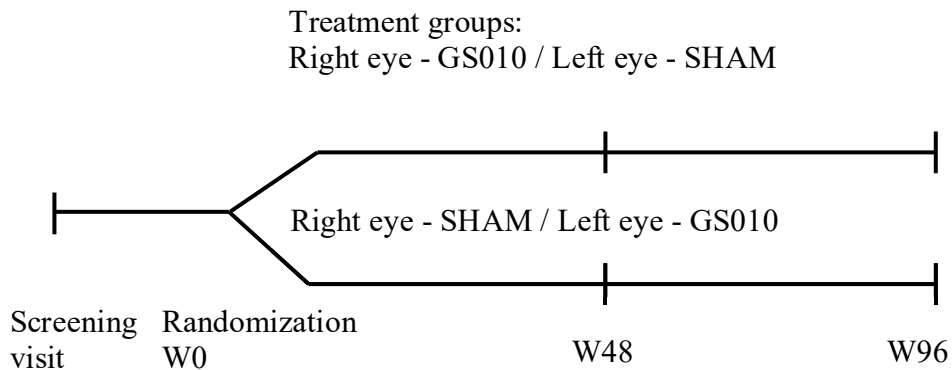
GS-LHON-CLIN-03B is a Phase III double-masked, sham-controlled, multicenter, multi-country study. Recruitment is competitive.

Both eyes will receive standard antiseptic preparation, administration of topical local ocular anesthetic agents and will undergo pupillary dilation. Administration of an intra-ocular pressure lowering agent will precede treatment. GS010 will be administered once during the study via a single intravitreal (IVT) injection. Sham IVT injection will be performed by applying pressure to the eye at the location of a typical procedure using the blunt end of a syringe without a needle.

The right eye of each subject will be randomly allocated to receive either GS010 or sham treatment in a 1:1 allocation ratio. The fellow (left) eye will receive the treatment not allocated to

the right eye. Therefore, each subject will receive GS010 in one eye and sham treatment in the fellow eye.

The schematic design of the study design is given below:



An algorithm with a testing hierarchy will be employed to determine the best- and worst-seeing eye of each subject prior to randomization. The randomization will be based on the right and left eye, because if the right eye is treated the left eye will receive sham and vice versa the treatment groups will be automatically balanced. However, the best- or worst-seeing eye at entry can be a confounding factor, and it cannot be discarded that the treatment effect will be the same in the best and worst seeing-eye. Secondarily, in order to verify this hypothesis with optimal power, an adaptive randomization technique called Efron's minimization ([Efron 1971](#)) method will be used to minimize the imbalance of treatment groups between the best-seeing eyes and worst-seeing eyes.

Masking will be accomplished with sham injection of the fellow eye. Thus treatment allocation will be double-masked for both the subject and the investigation (follow-up) team, except the physician and medical team that will perform the GS010 administration and sham injection and the first follow-up visit. Centralized analysis of the Early Treatment Diabetic Retinopathy Study (ETDRS) visual acuity results, automated visual fields, SD-OCT, color vision and contrast sensitivity will be performed.

## 3.2 DURATION AND SCHEDULE OF STUDY PARTICIPATION

### Study Schedule

The trial is divided into the following visits:

Screening Visit (Visit 1): From 28 to 2 days before the investigational medicinal product (IMP) administration. This visit will allow Investigators to assess subject eligibility based on selection and non-selection criteria.

Inclusion Visit (Visit 2): 1 day before the IMP administration. This visit will allow Investigators to confirm the subject's eligibility based on inclusion and exclusion criteria. Subjects will undergo baseline ophthalmological testing.

After the subject is confirmed to be eligible for the study, the right eye will be randomized to GS010 or sham treatment based on a predefined central randomization scheme.

Treatment Visit (Visit 3): Day of the IMP administration. Subjects will receive GS010 by IVT injection according to the randomized treatment assigned. The fellow eye will receive the alternative treatment.

Follow-up visits (Visits 4 to 12): Nine follow-up visits after the IVT procedure will be conducted at 24 hours and 2, 4, 8, 12, 24, 48, 72, and 96 weeks.

The study plan is shown in Section 4.2 of V3.0 of the protocol.

#### Study duration:

Initiation of the trial with the first subject's first visit is planned for late Q4 2015. The estimated recruitment period is 12 months and will be dependent on recruitment rate. Total study follow-up is 96 weeks and the study will end with the last subject's last visit. The last subject's last visit is estimated to occur late Q4 2018. Therefore, the estimated study duration is 3 years.

At the end of this initial study period, a long-term safety and efficacy study as detailed in a future separate protocol will be initiated for further evaluation of the safety and efficacy durability. The long term safety study duration will be conducted according to current regulatory authority recommendation.

### **3.3 SAMPLE SIZE CONSIDERATIONS**

Sample size calculation is based on the primary endpoint (LogMAR) and on the paired comparison of treated and sham eye. There is no published data regarding the within subject variance and the sample size calculation through paired comparison will therefore be very speculative for two reasons. The first reason is the within subject correlation between right and left eye and the second reason is the absence of information on the within and between subject variance of the LogMAR. Usually the within subject correlation is positive and the use of a null correlation is conservative. The absence of information on the LogMAR requires a second rough approximation of the mean and variance of LogMAR because the mean of ETDRS (Lam et al. 2014) is highly correlated to the variance of ETDRS ( $R^2 = 0.839$  for the log-log relationship). This correlation is in favor of the use of the LogMAR to stabilize the variance but the transformation from ETDRS mean and variance to LogMAR mean and variance requires additional assumptions. Starting with a mean 23.4 and a standard deviation of 28.3 (Lam et al.

2014) and a difference of 20 letters on average between treated and sham eye, using the relationship between the log mean and the log variance obtained with Lam et al. (2014) results ( $\log \text{variance} = 1.743 \log \text{mean} + 0.48$ ,  $R^2 = 0.839$ ) and assuming a perfect lognormal distribution of ETDRS, the difference in means of the lognormal distributions is 0.65 (3.32 - 2.67) and the pooled standard deviation of the log distributions is 0.914. Assuming no correlation between both eyes, the within subject standard deviation is  $(2 \times 0.914^2)^{0.5} = 1.29$ . The standardized difference in means is therefore  $0.65/1.29 = 0.504$  rounded to 0.50. The sample size required to get a power of 85% is exactly 36 subjects.

In GS-LHON-CLIN-01(Phase I/II study), at week 48 the standard deviation (SD) for visual acuity of the between-eye difference in change from baseline was calculated as 0.594. Considering that in Phase III studies of GS010 detection of a 15 ETDRS letters or 0.3 LogMAR difference between treated and untreated eyes at outcome is a clinically reliable difference, the SD of the difference of 0.594 would provide a power/probability of 84% to detect such a difference based on the enrolment goal of 36 subjects each in the RESCUE and REVERSE Phase III clinical trials.

### 3.4 RANDOMIZATION

The right eye of each subject will be randomly allocated to receive either GS010 or sham treatment in a 1:1 allocation ratio. The fellow (left) eye will receive the treatment not allocated to the right eye. Therefore, each subject will receive GS010 in one eye and sham treatment in the fellow eye. Prior to randomization, each subject's best- and worst-seeing eye will be determined. Efron's minimization method will be used in the randomization process to balance as much as possible the two treatment groups (GS010 and sham) in the best-seeing eyes.

To enroll a subject (Visit 1), the Investigator will call an interactive response system (IRS) and provide brief details about the subject to be enrolled. Each subject will receive a subject number assigned at screening, which will serve as the subject identifier throughout the study. The subject number will be required in all communication between the Investigator (or designee) and the IRS regarding a particular subject. Subject numbers will be tracked via the IRS.

To randomize a subject (Visit 2), the Pharmacist (or designee) will call the IRS and provide brief details about the subject to be randomized. Eligible subjects will be randomized in a 1:1 ratio (right/left eye), taking into account the Efron's minimization method, using the IRS.

The IRS User Guide will describe the steps of the randomization process from the subject's screening up to before IVT injection.

An algorithm with a testing hierarchy will be employed prior to randomization to determine the best- and worst-seeing eye of each subject, as follows:

1. ETDRS visual acuity score for subjects able to read the ETDRS chart. The eye with the higher ETDRS score is the best-seeing eye. For subjects who cannot read the ETDRS chart ("off chart" subjects) a standard study procedure will be used to determine visual acuity based



on counting fingers at a given distance. The eye with the best visual acuity based on this method is the best-seeing eye. If one eye of a subject is able to read letters on the ETDRS chart and one eye cannot, the eye able to read the ETDRS chart is the best-seeing eye.

2. SD-OCT parameters (to be used if there is no inter-eye difference based on Criterion 1):

- i) The initial SD-OCT parameter will be the total volume of the RGC layer of the macula. The best-seeing eye will have the greater volume and quadrant thickness, unless there is macula edema. A difference of  $\geq 5\%$  is considered significant to determine greater volume.
- ii) The second SD-OCT parameter (to be used if there is no inter-eye difference in criteria 2i) will be the combined volume of the RGC layer of the inner and outer nasal quadrants of the macula. The best-seeing eye will be the eye with greater volume. A difference of  $\geq 5\%$  is considered significant to determine greater volume.

3. Contrast sensitivity measured with the Pelli-Robson chart (to be used if there is no inter-eye difference in criteria 1 or 2). The eye with the best Log CS (contrast sensitivity) score is the best-seeing eye.

4. If the best-seeing eye is not possible to be determined based on any of the criteria in the algorithm above, the masked investigator will determine the best-seeing eye based on the totality of the Visit 1 vision testing results and subject testimony.

Perceptive Solution, an IRS company, will be responsible for the randomization code generation and actual randomization.

## **4. STUDY VARIABLES AND COVARIATES**

### **4.1 PRIMARY VARIABLE**

The primary endpoint will be the ETDRS visual acuity (quantitative score) at Week 48 after IVT injection. The subjects' LogMAR scores, which are derived from the number of letters they read on the ETDRS chart, will be used for statistical analysis purposes. LogMAR scores for subjects unable to read the ETDRS chart (i.e. off-chart subjects) will be obtained per the pre-defined Standard Operating Procedure and the protocol. The change from baseline in each eye will be the primary response of interest.

### **4.2 SECONDARY VARIABLES**

#### **4.2.1 Efficacy**

- Measure of parameters, including RGC layer, INL, ONL thickness/volume, RNFL thickness, and topographical maps of these anatomical regions, obtained with high resolution SD-OCT of the posterior pole and optic nerve at Week 48 and Week 96.

- ETDRS visual acuity (quantitative score) over the follow-up period and at Week 96 after IVT injection. Change from baseline of the LogMAR scores will be used for statistical analysis purposes.
- Response status to treatment at week 48 and 96 after IVT injection.
- Mean Deviation (MD) in decibels of sensitivity and Foveal Threshold Sensitivity will be used at Week 48 and Week 96.
- Measure of contrast sensitivity with the Pelli-Robson chart at Week 48 and Week 96.
- Measure of color vision with the Farnsworth-Munsell 100 Hue color vision test at Week 48 and Week 96.

#### 4.2.2 Safety

- Adverse events (AEs) and serious adverse events (SAEs), including those that are treatment-emergent and non-treatment-emergent, throughout the study period and at each study visit. Incidence and severity of systemic and local (ocular) AEs and SAEs will be determined at each clinical site and for the entire study cohort.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

### 4.3 PREDETERMINED COVARIATES AND PROGNOSTIC FACTORS

RGC layer, INL, ONL thickness/volume and topographical maps of these anatomical regions, best or worst-seeing eye (eye status) as determined prior to randomization and duration of vision loss at the time of treatment could be significant prognostic factors for the treatment effect. Sensitivity analysis will be performed for RGC layer thickness/volume and topographical map and eye status (best or worst-seeing eye) and duration of vision loss at the time of treatment by including them in the primary model as a covariate.

## 5. DEFINITIONS

### 5.1 STUDY DAY

All study days on or after the treatment administration will be calculated as date of assessment minus date of treatment administration +1. Study days before the treatment administration will be calculated as date of assessment minus date of treatment administration. In case of missing or incomplete dates no study days will be calculated.

### 5.2 BASELINE

The baseline value is defined as the last non-missing value prior to the start time of treatment administration (Visit 3). If assessment and treatment administration occur on same day and time is not captured, it will be considered as baseline. For OCT parameters baseline is defined as average of Visit 1 and Visit 2 assessments.

### 5.3 CHANGE FROM BASELINE

The change from baseline value is defined as the value at a given time point minus the baseline value.

### 5.4 DERIVED VARIABLES

#### ETDRS Score and LogMAR (on chart and off chart)

ETDRS score is calculated by the examiner and written on the source documents and entered in the eCRF.

LogMAR (on chart and off chart) will all be derived in eCRF from the Snellen/decimal acuity.

### LogCS

LogCS is obtained based on which lines the subject can read on the chart and the LogCS is then entered in the eCRF.

### Total error score

Total error score is derived by the software and then entered into the eCRF.

### Responder

#### Eye Responder

##### Definition 1

Responder at week 48 and also at week 96, will be defined by an improvement of at least 15 letters in the ETDRS score compared to baseline, *or* being greater than a Snellen acuity equivalent of 20/200. Early discontinuations (prior to W48 for the primary endpoint) will be considered as non responder. For week 96 two analyses will be performed: One for completers and one for subjects who don't have the Week 96 value available. If the week 96 ETDRS score is not available, then the previous visit value will be used.

##### Definition 2

Responder, at week 48 and also at week 96, will be defined by having an ETDRS Score equal to or greater than 20 letters. Early discontinuations (prior to W48 for the primary endpoint) will be considered as non responder. For week 96 two analyses will be performed: One for completers and one for subjects who don't have the Week 96 value available. If the week 96 visual acuity score is not available, then the previous visit value will be used.

#### Subject Responder

Responder, at week 48 and also at week 96, will be defined as a subject whose ETDRS score of the treated eye is at least 15 letters better than the sham eye *or* whose treated eye has a LogMAR acuity of at least 0.3 LogMAR better than the sham eye. Early discontinuations (prior to W48 for the primary endpoint) will be considered as non responder. For week 96 two analyses will be performed: One for completers and one for subjects who don't have the Week 96 value available. If the week 96 visual acuity score is not available, then the previous visit value will be used.

### Vision Loss Duration at The Time of Treatment

Vision loss duration in days will be based on the difference between the date of onset of vision loss per eye and screening date.

#### ETDRS/LogMAR and on-chart and off-chart:

- I. 1 ETDRS line = 5 letters
- II. 1 ETDRS line = 0.1 LogMAR
- III. 0.1 LogMAR = 5 ETDRS letters
- IV. 15 ETDRS letters = 0.3 LogMAR

V. A less positive/more negative LogMAR is an improvement of vision

VI. An increase in the number of ETDRS letters read is an improvement of vision

On-Chart subjects: defined as being able to read at least 3 ETDRS letters on a single line on the ETDRS chart at either 4 or 1 meters

- a. Have an obtainable Snellen equivalent obtained directly from ETDRS chart
- b. Each Snellen equivalent is assigned a LogMAR value

Off-chart Subjects: defined as a subject who cannot read at least 3 ETDRS letters on a single line of the ETDRS chart at either 4 or 1 meters and thus have no Snellen equivalent obtainable from ETDRS chart

- c. These subjects are categorized as able to either count fingers (CF), detect hand motion (HM), detect light perception (LP) or are no-light perception (NLP)
- d. For those categorized as either CF or HM a method for deriving LogMAR visual acuity is pre-specified
- e. The distance at which CF/HM vision was obtained is recorded per the standard operating procedure
- f. Study database derives the LogMAR based on this distance and the finger/hand dimensions of the examiner, per pre-specified method.

#### Treatment-Emergent Adverse Events (TEAE)

Treatment-emergent adverse events (TEAEs) are defined as AEs that started after study treatment administration at Visit 3 or that represent an exacerbation of a condition that is present at baseline after treatment administration. Worsening of visual acuity determined by the investigator to be due to progression of LHON will not be considered as an AE. Any AE with an unknown start date will be considered treatment-emergent if the event does not stop prior to study treatment administration. AEs with partial dates are described in [Section 5.5.2](#)

#### Prior and Concomitant Treatment

Prior treatments will be defined as any treatments with start date and time prior to the study treatment administration date and time.

Concomitant treatments will be defined as any treatments ongoing at the start of study treatment administration or with a start date and time on or after the study treatment administration date and time. Treatments with partial dates will be dealt similarly as AEs with partial dates described in [Section 5.5.2](#). If the start time is missing, then treatments will be assigned 'concomitant' based on start date of treatments only if it is on or after the treatment administration date.

[REDACTED]

[REDACTED]



## 5.5 HANDLING OF MISSING DATA AND/OR INVALID DATA AND OUTLIERS

### 5.5.1 Efficacy

In case of missing data at Week 48, the following rules will be applied in the primary imputation method:

Rule 1: If the LogMAR is available at the previous and the following visits then the imputed value at Week 48 will be a linear interpolation of LogMAR.

Rule 2: If there is no following visit due to drop out then the relative change ( $RC_i$ ) from visit  $V_i$  to visit  $V_{i+1}$  will be calculated at each visit from the baseline to the visit preceding Week 48 in the group treated with a sham. The last available visit ( $V_i$ ) with data ( $X_i$ ) will be used as the starting time point. The imputed score at  $V_{i+1}$  for each eye will be equal to  $X_{i+1} = X_i * RC_i$ . The imputed score at the next missing visit  $V_{i+2}$  will be equal to  $X_{i+2} = X_{i+1} * RC_{i+1}$ , and so on up to Week 48 visit. The  $RC_i$  will be the same for the treated and sham eyes.

With the primary method of imputation, the between eye difference at Week 48 will be proportional to the between eye difference at the time of withdrawal because the same RCI coefficients are used for both eyes.

The secondary imputation method leading to an additional sensitivity analysis will apply the same Rule 1 but Rule 2 will be based on the evolution of the difference between treated and sham eye. RCI is then the relative change from visit  $i$  to visit  $i+1$  of the difference between treated and sham eye in all subjects attending visit  $i$  and  $i+1$ . This method assumes that the evolution of drop-out is similar in terms of between eye differences to that of subjects who were followed-up.

Other:

When a test result score/value for the Pelli-Robson Contrast test, or the Farnsworth-Munsell 100 Hue Color test is missing at any visit, for the reason that the subject was unable to perform because the vision was too poor or unable to perform within the constraints of time limitation pre-defined in the Standard Operating Procedure (for the Farnsworth-Munsell 100 Hue Color test), the following convention will be utilized: the worst possible score for the given test will be utilized as the value for that visit.

### 5.5.2 Safety

#### Missing/Partial Start Date for AEs

The following rules will be applied to missing or incomplete start dates when determining if an AE is treatment emergent. The rules will be applied first before determining the treatment emergent status and imputed dates will be treated as complete dates for this purpose. These rules are intended to lead to a conservative assessment of treatment emergence. The imputed dates are only used for determining treatment emergence and the recorded partial dates will be displayed in listings. The same rules will be applied for prior and concomitant medications as per [section 5.4](#)

If the year of onset of the AE is missing, then the AE is considered treatment emergent.

If the year of the AE onset date is complete, but the month and day of the onset of the AE are missing, then the following rules apply:

- If the year of the AE is the same as the year of the study treatment administration then the AE is considered treatment emergent, and the AE onset date will be set to the study treatment administration date.
- If the year of the AE onset date is after the year of study treatment administration, then the date is set to 01 January and the AE is considered treatment emergent.
- If the year of the AE onset date is before the year of study treatment administration date, then the AE is not treatment emergent.

If the year and month of the AE onset date are complete, but the day of the onset of the AE is missing, then the following rules apply:

- If the year and month of the AE onset date is the same as the year and month of the treatment administration date, then the AE is considered treatment emergent and the AE onset date will be set to the treatment administration date.
- If the year and the month of the AE onset date are complete, and the month of the AE onset date is after the month of the study treatment administration date, then the day of the AE onset date will be set to the first day of the month. The AE will be assigned as treatment emergent.
- If the year and the month of the AE onset date are before the year and the month of treatment administration date, then the AE is not treatment emergent.

#### Missing/Partial End Date for AEs

If the end date is missing and the start date is on or after the date of treatment administration, then the AE is considered treatment emergent

#### Missing/partial dates for prior and concomitant treatments

Missing and partial dates of prior and concomitant treatments will be dealt similarly as partial dates for AEs.

If the year and month of the medication end date are complete and there is no day value, and the month of the medication end date is on or after the month of the study treatment administration date, then the day of the medication end date will be set to the first day of the month. The medication will be assigned as prior or concomitant medication based on the imputed date.

#### Missing severity or relationship to study treatment/study procedure of an AE

If the intensity of an AE is missing, an intensity of “severe” will be imputed.

If the relationship to study treatment/study procedure is missing, a relationship of “probable” will be imputed.

#### Missing/Partial End Date for Other Safety Variables

For other safety variables a similar approach for missing end dates will be followed in the data derivation of the Analysis Data Model (ADaM) specifications.

The same approach will be undertaken for the imputation of missing/partial end dates related to prior and concomitant treatments.

## **6. ANALYSIS SETS**

### **6.1 INTENTION-TO-TREAT**

The intent-to-treat population will consist of all subjects who are randomized and receive the actual study treatment (GS010). This population will be the population for the primary efficacy analysis.



## **6.2 PER-PROTOCOL POPULATION**

The per-protocol population (PP) will consist of all subjects who are randomized, receive the actual study treatment (GS010 and sham treatment) and complete the week 48 ETDRS assessment without any major protocol deviation. This population will be the population for the supportive analysis of primary efficacy analysis.

## **6.3 SAFETY**

The safety population is defined as those subjects who received study treatment (GS010). This population will be used as the population for all safety analyses.

## **7. INTERIM ANALYSES**

An interim analysis is not planned for this study.

### **7.1 TIMING OF PRIMARY ANALYSIS**

Primary endpoint analysis will be performed as soon as the last subject completes the Week 48 assessment and data are cleaned. At that point, all subjects will be recruited and followed for at least 48 weeks except in case of premature withdrawal. Many subjects will have reached Week 96 but not all of them. The study will be formally unmasked for the efficacy analysis at Week 48. The study will still continue up to scheduled end (Week 96). Although the study will have been unmasked for the study statistician, every effort (see Appendix V in Section 20.5 of the V3.0 protocol) will be invested to maintain the masking of subjects and personnel in charge of the conduct of the study including study managers, data managers, CRAs, investigators (except international coordinating investigators per Appendix V in Section 20.5 of the V3.0 of the protocol).

**Data Safety Monitoring Board (DSMB):** An independent DSMB will be constituted and responsible for periodically reviewing study data to assure the continued safe conduct of the study. The DSMB will meet to review the data at least every 6 months. Operational and logistical details will be provided in a separate DSMB charter.

## **8. DATA REVIEW**

### **8.1 DATA HANDLING AND TRANSFER**

Details regarding the data handling and data transfer process can be found in the PRA Clinical Informatics Plan.

### **8.2 DATA SCREENING**

Beyond the data screening built into the PRA Data Quality Plan, the PRA programming of analysis datasets, tables, figures, and listings (TFL) provides additional data screening. Presumed

data issues will be output into SAS logs identified by the word “Problem” and extracted from the logs by a SAS macro and sent to Data Management.

Review of a pre-freeze TFL run on clean subjects and a post-freeze TFL run on the frozen database allow for further data screening prior to lock. The post-freeze TFL at Week 48 will be discussed with GenSight in a masked data review meeting to identify any final data issues and seek corrections prior to data-base lock. The PRA statistician and GenSight must approve database lock. The same procedure will be performed before the W96 database lock.

## 9. STATISTICAL METHODS

All analyses will use SAS version 9.4 or higher.

Unless otherwise noted, categorical variables will be summarized using counts and percentages. Percentages will be rounded to one decimal place, except 100% and 0% which will be displayed without any decimal places.

Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), median, first quartile (Q1), third quartile (Q3), minimum and maximum. Minimum and maximum will be rounded to the precision of the original value.

Rounding rules are generally as follows:

- If the original value has 0 decimal places: mean, median, SD, Q1, Q3 and 95% confidence intervals (CIs) have 1 decimal place.
- If the original value has 1 decimal place: mean, median, SD, Q1, Q3 and CIs have 1 decimal place.
- If the original value has 2 or more decimal places: mean, median, Q1, Q3, CIs, and SD have the same decimal places as the original measure, up to a maximum of 3 decimal places.
- P-values will be presented as four decimal places.

Summary TFLs will be in landscape orientation and display standard headers containing the sponsor name (“GenSight Biologics”) as well as a standard footer containing the name of the program used to generate the display, its path, and a date/time stamp. Headers, footers, titles, and column headers for tables will be repeated on every page of each display; pages will be numbered “Page x of y.” Visit labels in tables and figures will show visit name. Summary TFLs containing no applicable data will display a line reading, “No applicable data were reported” or similar language. Dates and times will be presented using the DDMMYYYY HH:MM format. Summary TFLs will be produced using SAS<sup>®</sup> Output Delivery System (ODS) to generate rich text format (RTF) files and/or PDF files, if necessary. The default font will be 9-point Times New Roman, and minimum margins will be top 1”, bottom 1”, left 1”, right 1” (US letter). Standard indentation for subheadings is 2 spaces, and standard footnote indicators will be lowercase letters (e.g., a, b, c). Individual files will be named with the table, figure, or listing number plus the program name (with more descriptions related to output), concatenated with

underscores and using leading zeros as necessary for alphabetical sorting (e.g., T\_14\_1\_1\_disp). Listings will be sorted by location (Right/left), treatment and subject number if appropriate. For the listings where location and treatment are not appropriate, listings will be sorted by subject number only.

## **9.1 SUBJECT DISPOSITION**

The number and percentage of subjects, who were screened, randomized, treated in the study, who completed the study, and the number of subjects who completed the Week 48 visit will be presented, together with the number and percentage of subjects who withdrew from the study prematurely and a breakdown of the corresponding reasons for withdrawal. The withdrawal data will be presented overall, as well as presenting withdrawal before Week 48 and withdrawal after week 48. This data will also be listed according to the eCRF. The number and percentage of subjects in ITT, PP and safety population will also be presented.

Week 48 data will be analyzed for the ongoing studies. The listing per subject of frozen visits (up to W48) will be presented along with status 'ongoing' for unfrozen visits. Once the individual listing is presented, all the data up to week 48 will be frozen and not disclosed. No data beyond week 48 will appear in the listing up until the data of all visits up to Week 96 is frozen. A table summarizing number of subjects with frozen data at each visit will also be presented.

Number of subjects who missed visits among the subjects who attended a particular visit will be presented. In addition, number of missing data for visual acuity and important SD-OCT parameters will also be summarized.

## **9.2 PROTOCOL DEVIATIONS AND VIOLATIONS**

All major protocol deviations will be identified. A sensitivity analysis of the primary end point will be performed based on per-protocol analysis population. All the major protocol deviations will be listed. The decision regarding the major protocol deviations will be taken during the masked data review meeting. All the major protocol deviations will be identified and recorded in the Clinical Trial Management System (CTMS) of PRA.

## **9.3 DEMOGRAPHIC AND BASELINE CHARACTERISTICS**

Demographic and baseline characteristics (gender, age, weight, height) will be summarized using descriptive statistics. Qualitative variables (gender) will be summarized using frequencies while quantitative variables (age, weight, height) will be summarized using mean, SD, median, Q1, Q3, minimum and maximum.

Demographic and baseline characteristics will be summarized for the ITT and Safety Population. The analysis will be presented in two ways: Best-GS010 vs Best-Sham and Worst-GS010 vs Worst-Sham.

This data will also be listed based on the data listed in the eCRF.

Medical history data will be summarized by using counts and percentages and will be presented by system organ class and preferred term. Duration of vision loss at the time of treatment will be summarized by treatment group (All-GS010 vs All Sham, Best-GS010 vs Best-Sham and Worst-GS010 vs Worst-Sham). For the All-GS010 versus All-Sham comparison the safety population will be used for analysis, whereas for the Best-GS010 versus Best-Sham comparison and the Worst-GS010 versus Worst-Sham comparison the ITT population will be used for analysis of the duration of vision loss.

## **9.4 TREATMENTS**

### **9.4.1 Extent of Study Drug Exposure**

This study includes a single IVT injection in randomized eye and a single sham IVT injection in fellow eye on same day. Therefore, no summary table will be produced for the exposure data. eCRF data related to IMP administration will be listed.

### **9.4.2 Concomitant Treatments**

Prior and concomitant treatments will be coded using the World Health Organization dictionary (WHO Drug Dictionary Version 01Mar2015 DDE, B2). For each medical entity all Anatomical Therapeutic-Chemical (ATC) codes will be assigned according the therapeutic classes to which the treatment belongs. Coding of Treatments is detailed in the Coding Guidelines.

Prior and concomitant treatments will be summarized for the Safety population by number and percentage of subjects of WHO Drug Dictionary ATC levels (ATC1, ATC2) and preferred term. All prior and concomitant treatment data will be listed.

## **9.5 EFFICACY ANALYSES**

All efficacy analysis will be conducted on the ITT population and analyzed by actual treatment received. Sensitivity analysis will be performed based on a PP Population.

### **9.5.1 Primary Variable**

Primary variable will be analyzed by using two approaches:

For EMA, the change of LogMAR from baseline to Week 48 will be analyzed between all-GS010 treated eyes and all-Sham treated eyes.

The FDA has recommended comparing GS010 vs Sham in the Best-seeing eyes as a primary analysis. Therefore, for the FDA change of LogMAR from baseline to Week 48 will be analyzed between Best-GS010 treated eyes and Best-Sham treated eyes as the primary analysis.

The approach to analyze the primary variable is different for EMA and FDA. For EMA the analysis is based on intra-subject comparison of change of LogMAR from baseline to Week 48. For FDA, the analysis is based on inter-subject comparison of change of LogMAR from baseline.

The two approaches are stated below in detail:

#### Primary Analysis for EMA

The null and alternative hypotheses for the primary endpoint are defined below

$$H_0: \mu_{\text{All-GS010}} = \mu_{\text{All-Sham}}$$

$$H_1: \mu_{\text{All-GS010}} \neq \mu_{\text{All-Sham}}$$

Where  $\mu$  presents the mean change from baseline to Week 48 in LogMAR. “All” represents all eyes treated with GS010 or Sham

The primary analysis will be performed on the ITT population. The change from baseline to Week 48 in the LogMAR acuity will be the primary response. The change from baseline of all-GS010 eyes will be compared to the change from baseline of all-Sham eyes (intra-subject comparison) using a mixed model of analysis of covariance (ANCOVA) using the following terms (or covariate) in the model:

- Subject as random factor
- Eyes of the subject as random factor,
- Treatment and the
- Baseline LogMAR for each eye.

The analysis will be performed based on alpha of 0.05. The statistical significance in favor of GS010 against sham will provide evidence of the advantage of using GS010 for halting visual acuity loss or improving visual acuity in subjects with LHON with duration of vision loss of more than 6 months and up to 12 months. The difference in the mean change from baseline (adjusted mean difference) to Week 48 between the two treatment groups and associated 95% confidence interval will be reported.

The center will not be included in the analysis as the number of subjects recruited per site will be very small. However, stratification by geographical area will be performed as a sensitivity analysis. Neighboring countries can be pooled together to reach at least 9 subjects per geographical entity. The pooling will be performed during the masked review of data.

An additional model will add the treatment by geographical interaction for lack of generalizability of treatment effect purposes.

All the model diagnostics will be checked.

#### Secondary analysis of primary variable for EMA

Secondary Analysis of primary variable for EMA will be based on the methodology for primary analysis of primary variable for FDA stated below using ANCOVA. For secondary analysis change of LogMAR from baseline at Week 48 will be compared between best-seeing GS010 eyes and best-seeing Sham eyes (inter-subject comparison). Also for secondary analysis change of LogMAR from baseline in worst-seeing eyes treated with GS010 will be compared worst-seeing eyes treated with Sham.

#### Other Analysis

Summary statistics for LogMAR and ETDRS score and change from baseline will be presented by treatment group (All-GS010/All-Sham).

A graph presenting the mean of each treatment group at baseline and the mean at Week 48 will be presented. The graph will be useful to assess whether the difference in the change from baseline is mainly due to the difference at baseline or the difference at Week 48.

#### Sensitivity Analyses

The primary analysis with the same model will be repeated for the following lists.

- 1) Per protocol
- 2) Randomized subjects even if they were not treated and assigned treatment by randomization instead of actual treatment received
- 3) Secondary handling of missing data
- 4) Model with geographical area effect
- 5) Model excluding major outliers if any
- 6) Analysis of absolute Week 48 LogMAR values
- 7) Baseline GCL thickness/volume and topographical map
- 8) Best or worst-seeing eye (eye status) added as a binary covariate
- 9) Duration of vision loss at the time of treatment added as a covariate in the model

All these analyses will be presented in a forest plot displaying treatment effect, 95% confidence intervals and p-value. The analysis will also be presented in tabular format.

10) Non-parametric Analysis: Wilcoxon signed-rank test will be performed as a sensitivity analysis to compare between all-GS010 eyes and all-Sham eyes

#### Primary Analysis for FDA

The null and alternative hypotheses for the primary endpoint are defined below

$$H_0: \mu_{\text{best-GS010}} = \mu_{\text{best-Sham}}$$

$$H_1: \mu_{\text{best-GS010}} \neq \mu_{\text{best-Sham}}$$

Where  $\mu$  presents the mean change from baseline to Week 48 in LogMAR and “best” indicates best-seeing eyes.



The primary analysis will be performed on the ITT population. The change from baseline to Week 48 in the LogMAR acuity will be the primary response. The change from baseline of the best-GS010 eyes will be compared to the change from baseline of the best-Sham eyes (inter-subject comparison) using analysis of covariance (ANCOVA) including treatment as a main effect

- and baseline LogMAR

as a covariate in the model. The analysis will be performed based on alpha of 0.05. The statistical significance in favor of Best-GS010 eyes against Best-sham eyes will show evidence of the advantage of using GS010 in the best eye for halting visual acuity loss or improving visual acuity in subjects with LHON with a duration of vision loss of more than 6 months and up to 12 months. The difference in the mean change from baseline (adjusted mean difference) to Week 48 between the two treatment groups and associated 95% confidence interval will be reported.

The center will not be included in the analysis as the number of subjects recruited per site will be very small. However, stratification by geographical area will be performed as a sensitivity analysis. Neighboring regions will be pooled together to reach at least 9 subjects per geographical entity. The pooling will be performed during the masked review of data. An additional model will add the treatment by geographical area interaction to test for lack of generalizability of treatment effect.

The same analysis will be repeated for worst-GS010 eyes against worst-Sham eyes.

All the model diagnostics will be checked.

#### Secondary analysis of primary end point for FDA

Secondary Analysis of primary variable for FDA will be based on the methodology described above for primary analysis of primary variable for EMA. For secondary analysis, change of LogMAR from baseline at Week 48 will be compared between all-GS010 eyes and all-Sham eyes (intra-subject comparison). The difference in the mean change from baseline to Week 48 between the two treatment groups and associated 95% confidence interval will be reported.

#### Other Analysis

LogMAR and ETDRS score for each visit and change from baseline will be summarized by treatment group (Best-GS010/Best-Sham). LogMAR and ETDRS score for each visit and change from baseline will also be summarized by treatment group (Worst-GS010/Worst-Sham).

A graph presenting the mean of each treatment group at baseline and the mean at Week 48 will be presented. The graph will be useful to assess whether the difference in the change from baseline is mainly due to the difference at baseline or the difference at Week 48.

#### Sensitivity Analysis

Primary analysis with the same model will be repeated for the following lists.

- 1) Per protocol
- 2) Randomized subjects including those not treated
- 3) Secondary handling of missing data
- 4) Model with country effect
- 5) Model excluding major outliers if any (defined as a standardized endpoint greater than the absolute value of 4.0)
- 6) Analysis of absolute Week 48 LogMAR values, without covariate adjustment
- 7) Baseline GCL layer thickness/volume and topographical map
- 8) Duration of vision loss at the time of treatment as a covariate

All these analyses will be presented in a forest plot displaying treatment effect, 95% confidence intervals and p-value.

10) Non-parametric Analysis: Wilcoxon Rank-Sum test will be performed as a sensitivity analysis to compare between Best-GS010 eyes and Best-Sham eyes.

### **9.5.2 Methods for Handling Dropouts and Missing Data**

Method for dealing with missing data has been detailed in [Section 5.5.1](#)

### **9.5.3 Multiplicity**

All the type I error ( $\alpha=0.05$ ) will be spent for the primary end point analysis. All the other analysis will be supportive analysis of the primary end point. Therefore, no adjustment of multiplicity will be necessary.

### **9.5.4 Pooling of Sites**

No pooling of sites will be applied for the primary analysis. However, a stratification by geographical area will be performed as a sensitivity analysis. Neighboring countries can be pooled together to reach at least 9 subjects per geographical entity. The pooling will be performed during the masked review of data.

### **9.5.5 Secondary Variables**

#### GCL thickness/volume and topographical map

The change of RGC layer thickness/volume and topographical map from baseline to week 48 will be based on ANCOVA including the subject and the eyes of the subject as random factors, the treatment, and the baseline GCL thickness/volume as covariates in the model. The analysis will be performed based on alpha of 0.05. The difference in the mean change from baseline (adjusted mean difference) to Week 48 between the two treatment groups (all-GS010 vs all-Sham and best-GS010 vs best-Sham) and associated 95% confidence interval will be reported.



---

### RNFL temporal quadrant and papillo macular bundle

The change of RNFL temporal quadrant and papillo macular bundle from baseline to week 48 will be analyzed same way as GCL volume and topographical map.

### Change from baseline of the LogMAR scores at Week 96

The treatment estimate at Week 96 will be based on the same statistical approach as the primary endpoint analysis for EMA between all-GS010 and all-Sham and Best-GS010 vs Best-Sham.

### Time course of the response

Time course analysis will compare time course of the response up to Week 48 using the mixed model procedure. The change from baseline to Week 48 in the LogMAR will be the response. The change from baseline of all GS010 eyes will be compared to the change from baseline of all sham eyes using a mixed model of repeated measure (MMRM) including the subject and the eyes of the subject as random factors, the treatment, Visit, eye status at baseline (Best or worst), interaction between eye status at baseline and treatment, interaction between visit and treatment and baseline as covariates in the model. The estimate of treatment effect at various time points and its 95%CI (without adjustment for multiplicity) will be presented. Unstructured covariance structure will be fitted. In case of any convergence problem with this covariance structure other covariance structures will be used.

Three graphical presentations will be performed:

- 1) A graph with the time course of the response for the treated best eye, the sham best eye, the treated worst eye and the untreated worst eye.
- 2) A graph with the time course of the response of the treated eye (irrespective of the status at baseline) and the sham eye (irrespective of the response).
- 3) A graph with the time course of the response of the treatment effect (delta between treatment groups) for the best Seeing Eye at baseline and for the worst Seeing Eye at baseline.

In the graphs the pertinent p values including that of interaction between eye status and treatment, time effect, treatment effect, treatment by time interaction will be displayed.

Two additional analyses with the same model (excluding eye status and treatment interaction) for the best-seeing Eye and for the worst-seeing will be performed (sub-group analysis).

The same analysis will be continued for up to Week 96-time course.

All the model diagnostics will be investigated.

### Responder Analysis

#### Subject responders

The responder analyses at both Week 48 and Week 96 will be based on a randomization test comparing the observed rate of responders to the expected rate of responders under  $H_0$  (i.e. absence of treatment effect). Expected rate is defined as responder rate in the Sham group based on an improvement of at least 15 letters in the ETDRS score compared to baseline, or being greater than a Snellen acuity equivalent of 20/200.

#### Eye responders

The responder analyses at both Week 48 and Week 96 will be based on a McNemar test comparing all-GS010 eyes against all-Sham eyes.

#### Parameters of High Resolution SD-OCT

A central reading center will perform quality control, analysis and interpretation of all SD-OCT data. The following parameters will be summarized:

- a. Retinal nerve fiber layer (RNFL) thickness
  - i. Total average (360 degrees)
  - ii. Temporal quadrant
  - iii. Papillomacular bundle (PMB)
  - iv. Nasal quadrant
  - v. Inferior quadrant
  - vi. Superior quadrant
- b. Central retinal/macula ETDRS area thickness, volume and topographical maps
  - i. Retinal Ganglion Cell (RGC) layer
  - ii. Inner Nuclear Layer
  - iii. Outer Nuclear Layer
  - iv. Total retinal thickness/volume

The parameter values and change from baseline at Week 48 and Week 96 will be summarized by all-GS010 eyes and all-Sham eyes using summary statistics for continuous variables.

Mean graphs will be produced for all the parameters.

#### Measure of the standardized automated visual fields obtained with HVF Analyzer II.

Mean Deviation (MD) in decibels of sensitivity and Foveal threshold sensitivities at Week 48 and Week 96 and change from baseline to Week 48 and Week 96 will be summarized by all-GS010 eyes and all-Sham eyes using summary statistics for continuous variables. Forest plot will be presented displaying the magnitude of treatment effect and 95% confidence interval along with P-values. MMRM will be used to calculate the treatment effect and 95% confidence interval. A model similar to the one used for the time course of analysis for the LogMAR, will be used.

All the HVF parameter data will be listed.

---

Measure of contrast sensitivity (CS) with the Pelli-Robson chart

LogCS at Week 48 and Week 96 and change from baseline to Week 48 and Week 96 will be summarized by all-GS010 eyes and all-Sham eyes using summary statistics for continuous variables. Data will be listed according to the data in the eCRF.

Forest plot will be presented displaying the magnitude of treatment effect and 95% confidence interval along with P-values. MMRM will be used to calculate the treatment effect and 95% confidence interval. A model similar to the one used for the time course of analysis for the LogMAR, will be used.

Measure of color vision with the Farnsworth-Munsell 100 Hue color vision test

Total error score from HCT at Week 48 and Week 96 and change from baseline to Week 48 and Week 96 will be summarized by all-GS010 eyes and all-Sham eyes using summary statistics for continuous variables.

Forest plot will be presented displaying the magnitude of treatment effect and 95% confidence interval along with P-values. MMRM will be used to calculate the treatment effect and 95% confidence interval. A model similar to the one used for the time course of analysis for the LogMAR, will be used.

Data will be listed according to the data in the eCRF. Mean graphs will be presented for total error score from Farnsworth-Munsell data.

Refraction for BCVA

The refraction data will be listed according to the eCRF.

[REDACTED]

## 9.6 SAFETY ANALYSES

All safety analysis will be conducted on the Safety population and summarized by treatment actually received.

### 9.6.1 Adverse Events

All AEs coded using MedDRA dictionary (version 18.1).

Adverse events will be summarized in two different ways, systemic TEAEs and treatment-emergent ocular AEs. Systemic AE data will be summarized by all subjects. Treatment-emergent ocular AEs will be summarized by Sham eye only, GS010 eye only, both eyes, regardless of the eye

The following summaries will be produced for systemic TEAEs and treatment-emergent ocular AEs using counts and percent:

- Overall summary of AEs
- TEAEs, including the number of events reported
- TEAEs by relationship to study treatment
- TEAEs by relationship to study procedure
- TEAEs by maximum severity
- TEAEs leading to study discontinuation
- SAEs
- AEs leading to death

Listings of AEs, SAEs, ocular AEs and AEs leading to death will also be presented.

Each summary above will be summarized by system organ class and preferred term.

Note that counting will be by subject, not event and subjects are only counted once within each system organ class or preferred term. Subjects with multiple events within a particular system organ class or preferred term will be counted under the category of their closest drug-related event within that system organ class or preferred term for the summary tables related to study drug and procedure.

Subjects with multiple events within a particular body system or preferred term will be counted under the category of their most severe event within that body system or preferred term for the summary table of TEAE by maximum severity.

A summary of AEs leading to discontinuation will be provided, grouped by system organ class and preferred term.



All AEs recorded on the eCRF will be listed.

\_\_\_\_\_



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



Protocol no: GS-LHON-CLIN-03B

Statistical Analysis Plan  
12 March 2018/ Version 1.0

## 10. VALIDATION

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



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



## APPENDIX 1 GLOSSARY OF ABBREVIATIONS

Glossary of Abbreviations:	
AAV2	Adeno-associated Viral Vector, Serotype 2
ADAM	Analysis Data Model
AE	Adverse event
ANCOVA	Analysis of Covariance
ATC	Anatomic Therapeutic Classification
CP	Clinical Programmer
eCRF	Electronic Case Report Form
CS	Contrast Sensitivity
DMC	Data Monitoring Committee
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
EMA	European Medicines Agency
ETDRS	Early Treatment Diabetic Retinopathy Study
EU	European Union
GCL	Ganglion Cell Layer
HVF	Humphrey Visual Field
IMP	Investigational Medicinal Product
IOP	Intra-ocular Pressure
IRS	Interactive Response System
ITT	Intention-to-treat
IRS	Interactive Response System
IVT	Intravitreal
LHON	Leber Hereditary Optical Neuropathy
LogMAR	Log of the minimal angle of resolution
MD	Mean Deviation
MedDRA	Medical Dictionary for Regulatory Activities

ODS	Output Delivery System
PP	Per Protocol
Q1	First Quartile
Q3	Third Quartile
█	██████████
RC	Relative Change
RTF	Rich Text Format
SAP	Statistical Analysis Plan
SAE	Serious Adverse Event
SD	Standard Deviation
SD-OCT	Spectral-Domain Optical Coherence Tomography
█	██
TEAE	Treatment Emergent Adverse Event
TFL	Tables, Figures and Listing
US	United States
█	██
WHODDE	World Health Organization Drug Dictionary Enhanced





























