GenSight Biologics

DISCLOSURE: REDACTED STATISTICAL ANALYSIS PLAN VERSION 10.0

Title: Efficacy and Safety of Bilateral Intravitreal Injection of GS010: A Randomized, Double-Masked, Placebo-Controlled Trial in Subjects Affected with G11778A *ND4* Leber Hereditary Optic Neuropathy for Up to One Year

NCT Number: NCT03293524

Protocol Number: GS-LHON-CLIN-05

Statistical Analysis Plan Version 10.0, Approval Date: 10-Oct-2024

Certain information within this statistical analysis plan has been redacted (ie, specific content is masked irreversibly from view with a black bar) to protect either personally identifiable information or company confidential information.

This may include, but is not limited to, redaction of the following:

- Named persons or organizations associated with the study.
- Proprietary information, such as scales or coding systems which are considered confidential information under prior agreements with license holder.
- Other information as needed to protect confidentiality of GenSight Biologics or partners, personal information, or to otherwise protect the integrity of the clinical study.

STATISTICAL ANALYSIS PLAN

Efficacy and Safety of Bilateral Intravitreal Injection of GS010: A Randomized, Double-Masked, Placebo-Controlled Trial in Subjects Affected with G11778A ND4 Leber Hereditary Optic Neuropathy for Up to One Year

Drug	GS010 (rAAV2/2-ND4)
Study name	GS-LHON-CLIN-05
Phase of development	III
Sponsor	**GenSight Biologics** 74 rue du Faubourg Saint-Antoine 75012 Paris France
EudraCT Number	2017-002187-40
IND Number	16538
NCT Number	NCT03293525

General information

Protocol	GS-LHON-CLIN-05 V5.0 (23 December, 2019)
Related documents	• CRF version 6 • SAP V5.0 from Iqvia • ICH E9
Document Owner	eXYSTAT 4, rue Ernest Renan 92240 Malakoff [redacted]

Version History

Version number and date	Author	Main changes / Comments / Justification
1.0-3.0 04Sep2018	[redacted]	Not Applicable First Version
4.0 18Jul2019	[redacted]	Copied over from [redacted] version and updated as per IQVIA template
5.0 27Dec2019	[redacted]	Describe long term safety and efficacy analyses per version 5.0 of the protocol. Updated efficacy section to add more detail based on the protocol.
6.0 26OCT2020	[redacted]	Updated efficacy and safety sections and add of Natural history SAP in appendix 6.
7.0 14APR2021	[redacted]	Updated after dry-run.
8.0 03MAY2021	[redacted]	Modification of LogMAR calculation: The LogMAR values for CF and HM based on the Karanjia equation; will not be used for analyses due to the difficulties encountered in measuring the scale parameters on the clinical sites: The parameters measured for the calculation of the Karanjia LogMAR were not recorded accurately by the clinical centers:

		• Difficulties in the adherence to the parameters measurements described in the SOP • Errors in the conversion of the reading distances, finger and interdigital widths from US metric to EU metric system, leading to unreliable measurements. It is decided to only use the validated Lange LogMAR assignments to count fingers (CF) or hand motion (HM) visual acuity (C. Lange et al. Graefes Arch Clin Exp Ophthalmol (2009) 247:137-142)..
9.0 26OCT2021	[redacted]	1. Error in the SF-36 calculation, add 1 step: Transform composite scores PCS and MCS and 8 Z-score to **Normalized scores = 50 + (Health Domain Z-score × 10)** (see end of APPENDIX 4. Scoring rules for multi-item scales in SF-36 (Ware 2007) 2. Modification of prior / concomitant medication definitions: see Changes to analysis from protocol (7)
9.1 17FEB2023	[redacted]	1. Update of time windows definition to follow Data review decision for 1.5-year analysis. 2. Modification of data missing handle
9.2 17SEP2024	[redacted]	Update following Magali Taiel review before final analysis
9.3 09OCT2024	[redacted]	Update following DRM of 25SEP2024
10.0 10OCT2024	[redacted]	Final version before database lock

Approvals

Date of Data Review: 25SEP2024

Name	Position	Date	Signature
GENSIGHT			
[redacted]	[redacted]		
eXYSTAT			
[redacted]	[redacted]		

Table of Contents

Abbreviations

AAV2	Adeno-associated Virus Vector 2
AESI	Adverse Event of Special Interest
ATP	Adenosine Triphosphate
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AR(1)	Autoregressive Order 1
ARH(1)	Heterogeneous Autoregressive Order 1
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
AU	Arbitrary Units
AUC	Area Under the Curve
BCVA	Best Corrected Visual Acuity
BP	Blood Pressure
BUN	Blood Urea Nitrogen
cDNA	Complementary DNA
CF	Count Fingers
CFR	Code of Federal Regulations
CI	Confidence Interval
CLIN-01	GS-LHON-CLIN-01
CLIN-03A	GS-LHON-CLIN-03A RESCUE
CLIN-03B	GS-LHON-CLIN-03B REVERSE
CMO	Contract Manufacturing Organization
CRA	Clinical Research Associate
CRO	Clinical Research Organization
CRR	Clinically Relevant Recovery
CS	Contrast Sensitivity

CSH	Heterogeneous Compound Symmetry
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	Data Safety Monitoring Board
eCRF	Electronic Case Report Form
ELISA	Enzyme-Linked Immunosorbent Assay
ELIspot	Enzyme-Linked Immunospot Assay
EMA	European Medicines Agency
ENR	All Subjects Enrolled Set
ETDRS	Early Treatment Diabetic Retinopathy Study
EU	European Union
EOS	End of Study
FA	Fluorescein Angiography
FDA	Food and Drug Administration (United States)
GCL	Ganglion Cell Layer
GCP	Good Clinical Practice
γGT	Gamma-glutamyl transpeptidase
GLMM	Generalized Linear Mixed Models
GMO	Genetically Modified Organism
HIV	Human Immunodeficiency Virus
HM	Hand Motion
HVF	Humphrey Visual Field
IATA	International Air Transport Association
IC50	Half Maximal Inhibitory Concentration
ICF	Inform Consent Form
ICH	International Council for Harmonisation
IEC	Independent Ethics Committee
IgG	Immunoglobulin G
IFNγ	Interferon gamma
IMP	Investigational Medicinal Product
IMPD	Investigational Medicinal Product Dossier

IOP	Intraocular Pressure
IRB	Institutional Review Board
IRS	Interactive Response System
ITT	Intention-To-Treat
IVT	Intravitreal
LHON	Leber (or Leber's) Hereditary Optic Neuropathy
LMM	Linear Mixed effects Model
LOCF	Last Observation Carried Forward
LP	Light Perception
LogMAR	Log of the Minimal Angle of Resolution
LogCS	Log of Contrast Sensitivity
LS	Least Square
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention-To-Treat
mL	Milliliters
mg	Milligrams
MMRM	Mixed Model for Repeated Measurements
mtDNA	Mitochondrial DNA
n	Number of subjects with available data
Nab	Neutralizing Antibodies
ND4	NADH Dehydrogenase 4 Gene
NHP	Non-Human Primates
NLP	No Light Perception
P.O.	Per Os (i.e. orally)
PP	Per Protocol
PT	Preferred Term
Q1	First (lower) Quartile
Q3	Third (upper) Quartile
QoL	Quality of Life
qPCR	Quantitative Polymerase Chain Reaction

rAAV2	Recombinant Adeno-Associated Virus 2
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SD	Standard Deviation
SD-OCT	Spectral-Domain Optical Coherence Tomography
SF-36v2	36-Item Short Form Health Survey, Version 2
SOC	System Organ Class
SOP	Standard Operating Procedure
TARM1	Treatment arm 1
TARM2	Treatment arm 2
TEAE	Treatment Emergent Adverse Event
TOEP	Toeplitz
TOEPH	Heterogeneous Toeplitz
VFQ-25	Visual Functioning Questionaire-25
WHO	World Health Organization

1. Introduction

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for Protocol GS-LHON-CLIN-05. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

2. Study objectives

2.1. Primary objective

To assess the efficacy of intravitreal GS010 compared to placebo intravitreal injection in second-affected/not- yet-affected eyes, at 1.5 years post-treatment, analyzing the change from baseline of the best-corrected visual acuity (BCVA) reported with the Log of the Minimal Angle of Resolution (LogMAR), in NADH Dehydrogenase 4 Gene (ND4) Leber's Hereditary Optic Neuropathy (LHON) subjects with vision loss up to one year.

2.2. Secondary objectives

The efficacy objectives listed below will be analyzed for the second-affected/not-yet-affected eyes as the main eye of interest (designated the primary eye). Exploratory efficacy analyses will be conducted for the first affected eyes, but those analyses are secondary to any analysis on the primary eye, which is the second- affected/not-yet-affected eye.

1. To assess the safety and tolerability of bilateral and unilateral intravitreal injection of GS010.
2. To compare the time course of the BCVA LogMAR response in second-affected/not-yet-affected eyes treated with GS010 compared to placebo treatment.
3. To assess, at 1.5, 2, 3, 4 and 5 years post-treatment, the efficacy of intravitreal GS010 compared to placebo intravitreal injection in second-affected/not-yet-affected eyes, by determining the difference in the rate of responder eyes.
4. To compare the time course of the response in second-affected/not-yet-affected eyes treated with GS010 compared to placebo treatment and to estimate the magnitude of the treatment effect at 1.5, 2-, 3-, 4- and 5-years post-treatment, with parameters measured by:
    a. Spectral-Domain Optical Coherence Tomography (SD-OCT);
    b. Humphrey Visual Field (HVF) 30-2;
    c. Pelli-Robson Low Vision Contrast Sensitivity (CS).
5. To verify whether a difference exists at 1.5, 2, 3, 4 and 5 years post-treatment, between first affected eyes and second-affected/not-yet-affected eyes treated with GS010, using both an intra-subject and inter-subject analysis, with
    a. LogMAR BCVA
    b. Parameters measured by SD-OCT
    c. Parameters measured by HVF 30-2, and
    d. Parameters measured with Pelli Robson Low Vision Contrasts Sensitivity.

6. To assess, at 1.5, 2, 3, 4 and 5 years post-treatment, the rate of responders in first affected eyes treated with GS010.
7. To assess humoral and cellular immune responses to Adeno-Associated Virus vector 2 (AAV2) after unilateral and bilateral intravitreal administration.
8. To assess the impact of bilateral intravitreal GS010 administration on Quality of Life (QoL) scales at 1.5, 2-, 3-, 4- and 5-years post-treatment.

3. Study design

GS-LHON-CLIN-05 is a Phase III, global, multicenter, randomized, double-masked for the primary analysis, placebo-controlled, clinical study. As LHON is a neurodegenerative disease, the goal is to administer GS010 as soon as possible upon confirmation of the LHON diagnosis and the causative mutation.

All subjects will receive a peri-treatment, systemic immune modulating regimen for the prevention or diminution of ocular inflammation related to intravitreal injection of the Investigational Medicinal Product (IMP).

3.1. General description

At the Screening Visit (Visit 1), the duration of vision loss of each eye of each subject will be determined and documented. The eye with the longest duration of vision loss will be the first affected eye and the eye with the shortest duration of vision loss will be the second affected eye. Eyes that have not yet experienced clinically manifested decline in visual acuity will have a value of zero days for vision loss duration and will be considered "not yet affected" eyes. As all subjects are required to have clinically manifested vision loss due to ND4 LHON to any extent in at least one eye, each subject will thus have one eye designated as the "first affected" eye and one eye designated as the "second-affected/not-yet-affected" eye. Subjects may report simultaneous onset of vision loss in both of their eyes, indicating true simultaneous onset of vision loss or inability of the subject to distinguish which is the first affected eye and which is the second affected eye. For these subjects, vision loss duration will be equal for both eyes.

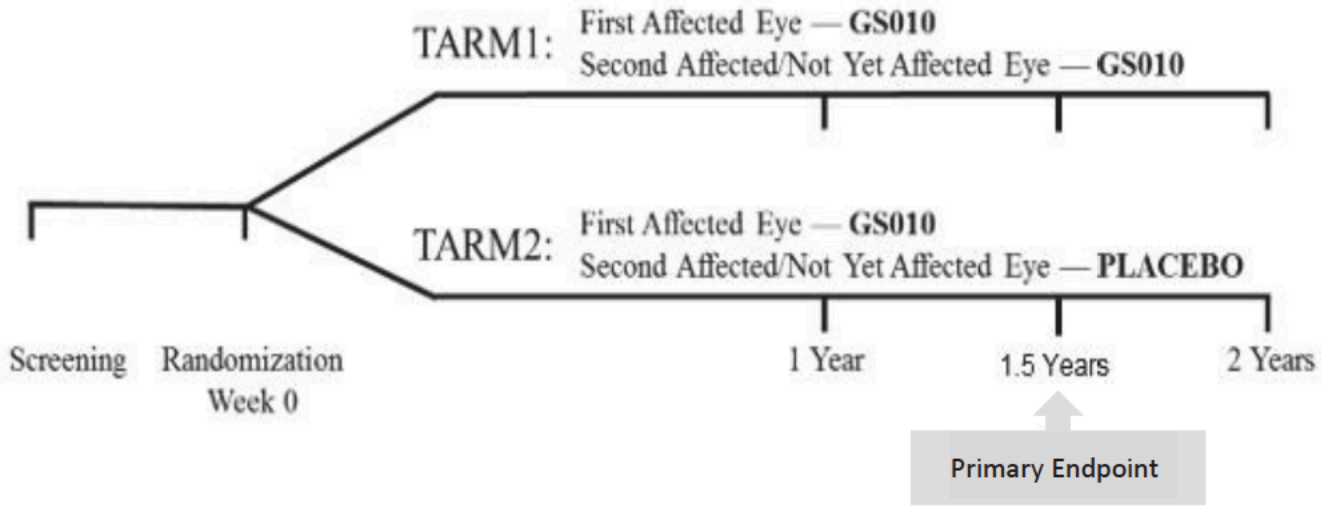
At the Inclusion Visit (Visit 2), new onset vision loss occurring since the Screening Visit (Visit 1) in an eye previously unaffected by clinically manifested visual acuity loss will be documented and vision loss duration will be recorded. The best-seeing and worst-seeing eye of each subject will be determined, based on the baseline vision testing performed at Visit 2. A pre-defined algorithm detailed in section 13.10.1 of the protocol to determine the best- and worst-seeing eyes will be employed.

An Interactive Response System (IRS) will assign a unique subject identification number at Screening Visit (Visit 1), which will be required for all communication between the Investigator (or designee) and the IRS. Subject numbers will be tracked via the IRS. The IRS will be used to enroll and randomize subjects based on a pre-defined central randomization scheme. An IRS User Guide will describe all steps for enrollment and randomization of subjects.

Subjects meeting all eligibility criteria for selection and inclusion will be randomized to treatment arm 1 (TARM1) or treatment arm 2 (TARM2) in a 1:1 allocation (see Figure 1). Subjects in TARM1 will receive intravitreal GS010 in their first affected eye and their second-affected/not-yet-affected eye. Subjects in TARM will receive intravitreal GS010 in their first affected eye and placebo intravitreal injection in their second- affected/not-yet-affected eye. For subjects who report simultaneous onset of vision loss in both eyes, the right or left eye will be selected randomly to serve as the second-affected/not-yet-affected eye.

If the right eye is selected randomly as the second-affected/not-yet-affected eye, the left eye will be designated as the first affected eye, and vice versa.

Figure 1: Study Treatment Arms; Post-Treatment Follow-up Period 1: up to 2 years post-treatment



Administration of an Intra ocular pressure (IOP) lowering agent of the investigator's choice will precede all intravitreal injections. Pre-intravitreal injection procedural preparation will include pupil dilation, topical antisepsis, and topical anesthesia. The IMPs (GS010 and placebo) will be administered with a standard intravitreal injection procedure. Each eye will undergo administration of the allocated treatment with a single intravitreal injection, performed as a separate procedure. Please refer to the Intravitreal GS010/Placebo Injection Guide for full details.

Per the Investigator's discretion and according to local standard accepted medical practices, treatment administration to both eyes may be performed at a single Treatment Visit (Visit 3) (i.e. intravitreal injection of both eyes of a given subject is performed the same day as two separate procedures) or may be performed on two consecutive days (i.e. one eye of the subject will be administered the allocated treatment one day and the next day the fellow eye of the subject will be administered the allocated treatment).

Inclusion Visit (Visit 2) vision test results will serve as the baseline values for both eyes of a given subject when the treatments are administered on the same day and when they are administered on consecutive days. Study Day 0 is the day that treatment administration is completed to both eyes; in the setting of treatment administration on consecutive days, this will be the second treatment day. The timing of visits and visit

windows for post-treatment follow-up visits will be determined from Study Day 0.

Adherence to the following rules is a requirement:

i. Administration of the allocated treatment to both eyes of a given subject must be completed within two consecutive days after the day of the Inclusion Visit (Visit 2).

ii. Follow-up Visit 4 must be performed 1-day following the completion of treatment administration to both eyes.

iii. The initial treatment administration must be performed the day following the Inclusion Visit (Visit 2). When both eyes are treated the same day, treatment must be performed the day after the Inclusion Visit (Visit 2).

Table 1 describes the two options for treatment administration timing.

TABLE 1: TREATMENT ADMINISTRATION OPTIONS

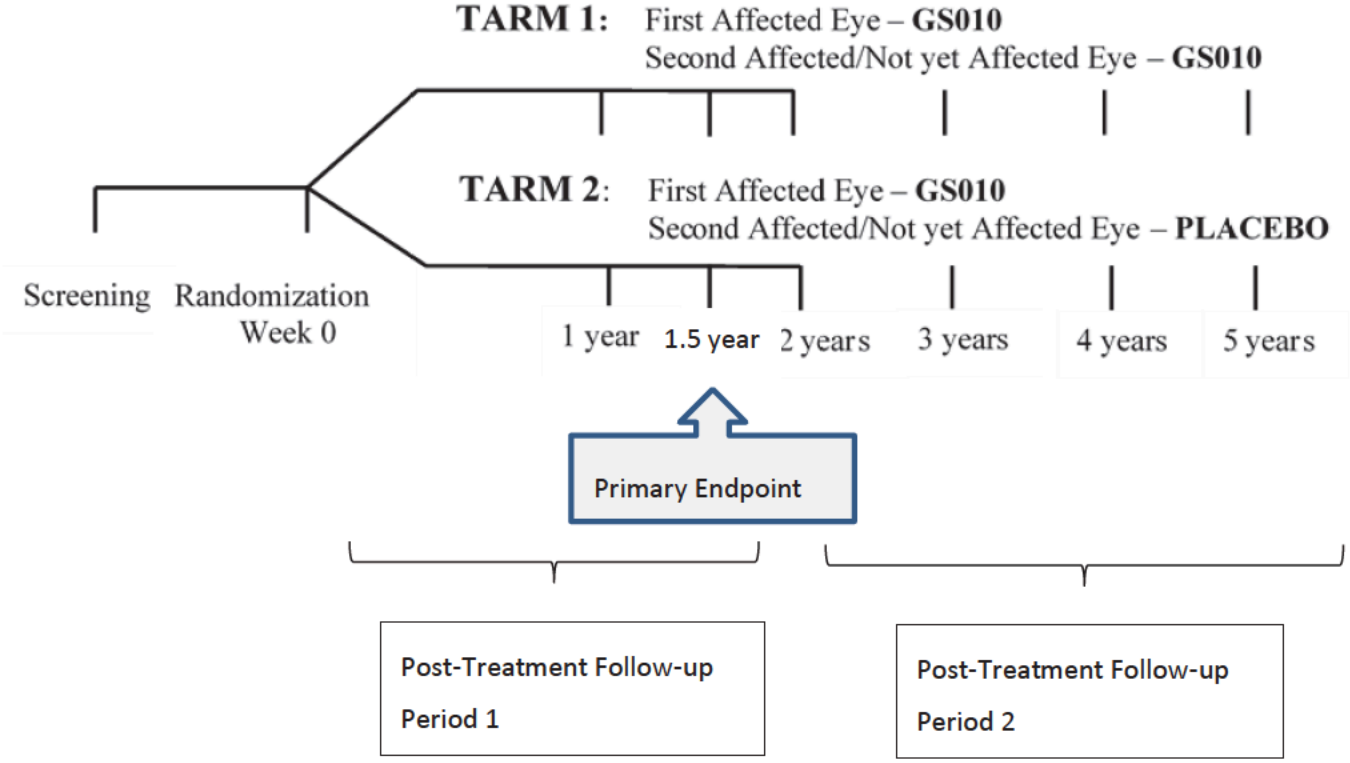
OPTION 1 (Eyes are treated on 2 consecutive days)		
Day -2	Day -1	Day 0
Inclusion	First IVT injection in one eye	Second IVT injection in fellow eye

OPTION 2 (Eyes are treated on the same day)	
Day -1	Day 0
Inclusion	Two IVT injections (one in each eye)

Masking will be accomplished with use of intravitreal injection of placebo. The pharmacy team will be the only study personnel/team not masked to treatment allocation (for the primary analysis). The treating and follow-up physicians and respective study teams will thus be masked to treatment allocation for second- affected/not-yet-affected eye (i.e. the primary eye). Subject's treatment allocation will be double-masked.

Each subject will be invited to participate in the Long-Term safety and efficacy follow-up (Post-Treatment Follow-up Period 2). Preferably, the informed consent of the Post-Treatment Follow-up Period 2 will be provided by visit 13.

Figure 2: Study Treatment Arms; Post-Treatment Follow-up Periods 1 and 2: up to 5 years post-treatment



3.2. Sample size consideration

Sample size calculation is based on the primary endpoint which is the change from baseline to 1.5 years post- treatment of the visual acuity (LogMAR BCVA). Masked preliminary data of the Phase III RESCUE and REVERSE sham-controlled studies of GS010 and monitored data from the Phase I/IIa safety and tolerability study of GS010 are used to estimate the standard deviation of the change from baseline to each available visit. These data do not take into account a potential treatment effect in the RESCUE and REVERSE studies and if there is an effect, the standard deviation may be overestimated.

From this masked preliminary analysis, a standard deviation of 0.55 LogMAR has been estimated for the LogMAR change from baseline; this is reduced to 0.50 with the inclusion of baseline LogMAR as a covariate. Using this reduced value for Standard Deviation (SD) and a treatment effect of 0.3 LogMAR in the change from baseline, the requested sample size to reach a power of 80% (alpha = 0.05 two-sided) according to a classical Analysis of Covariance (ANCOVA) model including the baseline LogMAR value as covariate has been estimated using SAS PROC POWER to be 45 evaluable subjects per treatment group (i.e. 90 subjects in total). Depending on the number of subjects who terminate early, additional subjects may be enrolled through the end of the recruitment period.

3.3. Schedule of events

A table showing details of the assessments to be performed by scheduled visit is given in section 9.2 of the protocol. These are summarized by scheduled visit in section 4.2.1 of this document and in section 9.1.1 of the study protocol.

3.3.1. Study schedule

The study is divided into 2 successive periods:

1. The first period covers enrollment, treatment and initial Follow-Up Period 1 of subjects up to 2 years Post-Treatment, divided in 13 visits:
    a. Screening Visit (Visit 1)
    b. Inclusion Visit (Visit 2)
    c. Treatment Visit(s) (Visit(s) 3)
    d. Post-Treatment Follow-Up Period 1 (Visits 4 through 13), with 2 optional telephone visits
2. The second period, Post-Treatment Follow-Up 2, covers a period from 2 years to 5 years post- treatment, divided in 3 additional visits (Visits 14, 15, and 16).

Screening Visit (Visit 1)

Will occur from 28 to 7 days before the Treatment Visit(s) (Visit(s) 3). Informed consent signature will be obtained. Eligibility for selection will be assessed. Demographic and medical history will be obtained, and general systemic evaluations (e.g. general physical examination, vital signs, 12-lead ECG) will be performed for screening and to obtain baseline values for comparison to post-treatment values. Laboratory assessments will be obtained. Vision loss duration of each eye will be determined, and vision testing will be performed per the study schedule.

Inclusion Visit (Visit 2)

Will occur one day prior to the initial IMP/placebo administration. Eligibility will be confirmed based on the inclusion/exclusion criteria. Baseline visual evaluations and ophthalmological examinations will be performed for comparison to post-treatment visits. Baseline QoL scale assessment will be performed for comparison to post-treatment values. Best- and worst-seeing eyes will be determined prior to randomization. Duration of vision loss in the second affected eye will be documented if vision loss occurred after the Screening Visit (Visit 1).

Upon confirmation of eligibility for inclusion, subjects will be randomized to the treatment arms based on a pre-defined central randomization scheme.

Treatment Visit(s) (Visit(s) 3)

Subjects will receive the treatment as allocated based on the randomization. Per the Investigator's discretion and according to local standard accepted medical practices, treatment administration to both eyes may be performed at a single Treatment Visit 3 (i.e. intravitreal injection of both eyes of a given subject is performed the same day), or may be performed on two consecutive days (i.e. one eye of the subject will be administered the allocated treatment one day and the next day the fellow eye of the subject will be administered the allocated treatment).

Post-Treatment Follow Up Period 1: Visits 4 through 13, with 2 optional telephone visits

Subjects will attend ten follow-up visits conducted at Day 1, Week 2, Months 1, 2, 3, 6, and 9, and Years 1, 1.5, and 2 post-treatment administration (in reference to Study Day 0).

Post-Treatment Follow-Up Period 2: Visits 14, 15 and 16

Subjects will attend 3 additional follow-up visits conducted at Years 3, 4 and 5 post-treatment administration (in reference to Study Day 0).

3.3.2. Study Duration

Initiation of the trial with the first subject's first visit occurred in March 2018. The Post-Treatment Follow-up Period 1 will end with the last subject's last visit (Visit 13 = Year 2) in Q3 2021. The study is expected to be completed with the last subject's last visit at the end of the Post-Treatment Follow-up Period 2 (Visit 16 = Year 5) estimated to occur in Q3 2024. The estimated total duration of the study is six and a half years, including the time required to complete recruitment and all study visits for all subjects.

3.4. Changes to analysis from protocol

1. To address the impact of COVID-19 on study integrity, an extension of the protocol-defined window of time for performing the endpoint has been decided during the data-review before the database lock. Data collected before or after 1.5 year was assessed by the data review committee as relevant and representative of the missing 1.5-year data. Time windows are specified in SAP section 5.3.
2. In the protocol (section 16.4.1.14): In case of missing data up to and including 1.5 years in the primary (second affected/not yet affected) eye, multiple imputation will be used as the primary method of imputation for the following endpoints: BCVA measured in LogMAR, Pelli-Robson CS, and HVF 30-2.

   Approach defined in protocol section changed to approach in SAP section 13.1.3: Last observation carried forward method (LOCF) will be used of imputing missing data. In this disease, outside of the very initial phase, BCVA worsen slowly over time (natural history of the disease), whereas active treatment improves BCVA over time, LOCF will therefore tend to favor the placebo eye group and disadvantage the treated eye group.
3. In the protocol (section 16.4.1.12): Sensitivity Analysis: For ITT and mITT analyses, missing data will be imputed using multiple imputations.

   Modify by (section 6.4.1.14): Sensitivity analyses of the primary endpoint will be conducted only on observed data (data without imputed values) up to 1.5 years using a MMRM model.
4. Removal of the exploratory analysis coverage (section 16.4.2.4.3) because this analysis is similar to the MMRM model analysis.
5. Removal intra-subject analyses of difference at 1.5, 2, 3, 4 and 5 years post-treatment, between first affected eyes and second-affected/not-yet-affected eyes treated with GS010 (part of objective 5), based on FDA recommendation: "Although use of the contralateral eye to which the GT product is not administered as a control may potentially be considered, it is generally not recommended" Protocol section and analyses described that are not done in SAP:

13.1.6 – Duration of disease is no longer used

13.1.5 – Change of supportive analyses

13.2.3– No more time to event analyses

13.2.3– AUC is no longer done

6. Planned analyses

The following analyses will be performed for this study:

- Analyses for Data Safety Monitoring Board (DSMB) meetings
- Primary efficacy analysis at 1.5 years
- Post-Treatment Follow-Up Period 2 Analyses at 3 and 4 years
- Final Analysis at 5 years

7. Definition of prior and concomitant medication

Prior medication:

- In protocol: Prior medications are defined as prescription and non-prescription medications and preparations, including health and/or dietary supplements, **prior to signing informed consent**.
- Change from protocol: Prior medications are defined as prescription and non-prescription medications and preparations, including health and/or dietary supplements, **prior to first IVT**

Concomitant medication:

- In protocol: Concomitant medications and therapies are defined as prescription and non-prescription medications and preparations, health and/or dietary supplements other than randomized study treatment that the subject receives during the course of the study (i.e. any time after **signing informed consent**). This includes any medication started prior **to signing informed consent** and continued after **signing informed consent**.
- Change from protocol: Concomitant medications and therapies are defined as prescription and non-prescription medications and preparations, health and/or dietary supplements other than randomized study treatment that the subject receives during the course of the study (i.e. any time after **first IVT**). This includes any medication started prior **to first IVT** and continued after **first IVT**.

3.5. Data safety monitoring board

An independent DSMB will be convened for data and safety review at least every 6 months until all subjects have completed Post-Treatment Follow-Up Period 1, to assure the continued safe conduct of the study.

The DSMB may recommend the discontinuation of the study for safety reasons. Operational and logistical details will be provided in a separate DSMB charter.

3.6. Primary efficacy analysis

Although there is not a planned interim analysis of efficacy, as described in the protocol section 16.4.1.15, the study will be formally unmasked for the primary efficacy analysis at 1.5 years (Week 78) following Sponsor Authorization of this SAP and interim Database Lock to allow registration if results meet the expectations. Efficacy and safety analyses will be included.

3.7. Post-treatment follow-Up period analyses

All planned analyses identified in this SAP will be performed by IQVIA Biostatistics following an interim database lock. The final analyses for the Post-Treatment Follow-Up Period 1 will be performed when all subjects have completed Visit 13, 2 years post dose. Three additional analyses will be performed when all subjects have completed Visit 14 (3 years post dose), Visit 15 (4 years post dose), and Visit 16 (5 years post dose). All efficacy analyses performed when all subjects have completed Visits 13, 14, 15, and 16 (i.e. 2-, 3-, 4-, and 5-years post dose) are considered secondary and will be supportive of the primary efficacy analysis.

4. Analysis populations

Agreement and authorization of subjects' inclusion/exclusion from each analysis set will be conducted prior to the unmasking of the study.

4.1. All subjects enrolled set

The all-subjects-enrolled (ENR) set will contain all subjects who provided informed consent for this study.

4.2. Intent-to-Treat population

The intent-to-treat analysis (ITT) population will consist of all randomized subjects. The analyses will be based on the planned treatment (as randomized). This ITT population is the primary efficacy analysis population. All missing endpoint data will be imputed for analyses using the ITT population.

4.3. Modified Intent-to-Treat population

The modified intent-to-treat population (mITT) will consist of all subjects who are randomized and receive study treatment (GS010 or placebo) in the second affected /not yet affected eye. In the case of an error in treatment allocation, the randomized treatment will be used in the analyses. The mITT population is supportive of the ITT analysis population. If all randomized subjects are treated, then this population will not be employed. All missing endpoint data will be imputed for analyses using the mITT population.

4.4. Per Protocol population

The Per Protocol (PP) population will consist of all subjects who are randomized and received study treatment (GS010 or placebo treatment) in the second affected /not yet affected eye. In the case of an error in treatment allocation, i.e. GS010 is administered to the eye allocated to receive placebo (or vice versa), the actual treatment received will be used in the analyses. Subjects identified with major protocol deviations, which are judged to potentially interfere with the interpretation of the treatment effect, will be removed from the PP analysis population. The PP population is supportive of the ITT analysis population. Major protocol deviations are identified and validated by the Sponsor; these will be detailed in a separate document. Missing data will not be imputed.

4.5. Per Protocol-5Y population (PP5Y)

The Per Protocol (PP) population at 5 years will consist of all subjects in the Per Protocol population who did not receive Idebenone after Year 1.5 visit. This population is supportive of the ITT analysis population for the 5Y analyses.

4.6. Safety analysis population

The safety population is defined as those subjects who received study drug (GS010 or Placebo) in at least one eye. Subjects will be classified according to treatment actually received. If there is any doubt whether a subject was treated or not, the subject will be assumed treated for the purposes of analysis.

This population will be used as the population for all safety analyses.

4.7. Follow-up population

The follow-up population is defined as subjects who completed Visit 13 and enrolled in the Post-Treatment Follow-Up Period 2. This population will be used to examine both safety and efficacy up to 5 years post treatment during the Post-Treatment Follow-Up Period 2.

5. General considerations

5.1. Reference start date and study day

Study Day will be calculated from the reference start date and will be used to show start/stop day of assessments and events.

Reference start date is defined as the date that treatment administration is completed in both eyes (Day 0).

Study Day = (date of event – reference start date).

In the situation where the event date is partial or missing, See section 6.2

5.2. Baseline

Unless otherwise specified, the baseline value is defined as the last non-missing value prior to the start day of the IVT injection.

5.3. Definitions of time points

The different scheduled visits are described in Section 3.3.1 and time windows around time points are defined as follow in the protocol.

Screening Visit (Visit 1) will be performed 7 to 28 days before Study Day 0. Inclusion Visit (Visit 2) will be performed 1 or 2 days before Study Day 0. Treatment Visit (Visit 3) will be performed on Study Days -1 (if applicable) and 0. Post-Treatment Follow-Up Period 1 visits will be performed as follows:

- Visit 4: Study Day 1
- Visit 5: Study Day 14 (±2 days)
- Visit 6: Study Day 28 (±3 days)
- Visit 7: Study Day 56 (±6 days)
- Visit 8: Study Day 84 (±9 days)
- Visit 9: Study Day 168 (±17 days)
- Visit 10: Study Day 252 (±17 days)
- Visit 11: Study Day 365 (±30 days)

- Visit 12 (Year 1.5 Visit): Study Day 548 (±30 days)
- End of Study (EOS) Visit 13 (Year 2 Visit): Study Day 730 (±30 days)

Post-Treatment Follow-Up Period 2 visits will be performed as follows:

- Visit 14 (Year 3 Visit): Study Day 1095 (±30 days)
- Visit 15 (Year 4 Visit): Study Day 1460 (±30 days)
- EOS Visit 16 (Year 5 Visit): Study Day 1825 (±60 days)

To address the impact of COVID-19 on study integrity, an extension of the protocol-defined window of time for performing the endpoint was decided during the data-review before the database lock. For this disease, remote measures remain informative.

After review during the BDR meeting and with the medical expert, the observed deviations in the performed visits were not considered as having a significant impact on the validity of the data.

Consequently, the visit windows are adjusted as follows:

Visit (days)	Lower time window	Timepoint	Higher time window
Visit 5	2	14	20
Visit 6	21	28	41
Visit 7	42	56	69
Visit 8	70	84	125
Visit 9	126	168	210
Visit 10	211	252	308
Visit 11	309	365	517
Visit 12	518	548	699

For analyses at 2 years, analyses at 3 years and analyses at 4 years and final analysis at 5 years, to take into account new visits done after the year 1.5 analyses, following time windows will be considered for visits 13, 14, 15 and 16:

Visit (days)	Lower time window	Timepoint	Higher time window
Visit 13	700	730	912
Visit 14	913	1095	1274
Visit 15	1275	1460	1638
Visit 16	1639	1825	2100

5.4. Retests, unscheduled visits and early termination data

Scheduled visits, unscheduled visits, and early discontinuation visits will be used independently of the nominal visit.

5.5. Windowing conventions

As stated in section 5.1, the timing of visits for post-treatment follow-up visits will be determined from Study Day 0.

All visits originally planned in the protocol schedule of events have a visit window that was specified during data-review (see section 5.3). In the case of multiple records for a subject within a particular visit window, then the assessment that is closest to the target day will be used. In the event that both (or all) the records are equidistant from the target day then the latest of the equidistant records will be used.

5.6. Statistical testing

The default significance level will be (5%); 95% confidence intervals will be produced, and all tests will be two- sided, unless otherwise specified in the description of the analyses.

5.7. Common calculations

For quantitative measurements, change from baseline at visit X will be calculated as:

Test Value at Visit X – Baseline Value

5.8. Software version

The statistical analysis will be performed using the SAS® software (Version 9.4 or higher).

6. Statistical considerations

6.1. Multicenter studies

This study will be conducted by multiple investigators at multiple sites internationally. Sites will be pooled per region, i.e. by continent of Asia, Europe and the USA.

The effect of region in the association between treatment and primary endpoint will be examined (see section 13.1.5)
.

6.2. Missing data

For incomplete dates following convention will be used:

- Birthdate missing: if birthdate is missing but an age has been recorded in the CRF, the year of the birthdate will be imputed as the following:

  Year of birth = year of screening – age at screening

  The birthdate will then be imputed with this year and 1st of January as day and month of birth.

- Date of onset of vision loss for each eye:
  - For missing day: '01' will be imputed (if month is not missing).
  - For missing start month: 'January' will be imputed (if start year is not missing).
  - For affected eyes (i.e. excluding non-affected eyes at baseline), a missing date of onset of vision loss for one eye, will be replaced by the date of onset of vision loss of the contralateral eye.
- For AEs and CMs incompletes dates, see Appendix 2
- For other incompletes dates, no imputation will be performed.

Missing efficacy data will only be imputed as described in sections 13.1.3 and 13.2.2.

6.3. Multiple comparisons/ multiplicity

All the type I error (alpha=0.05) will be spent for the primary endpoint analysis. Other analyses, including those performed at 2 years, described in section 15, will be performed as either secondary, sensitivity, supportive, or exploratory analyses with no multiplicity adjustment.

6.4. Examination of subgroups

Subgroups analyses will be performed. The following subgroups will be assessed for the primary endpoints (BCVA):

- Eyes on-chart for LogMAR BCVA at baseline,
- Eyes off-chart for LogMAR BCVA at baseline
- Pediatric (<18 years-old)/Adult populations,

- Males/Females,
- Previous idebenone treatment yes/no.

6.5. Covariance of subgroups

For the MMRM model, LMM and GLMM, autoregressive (AR(1)) covariance structure will be used to model the within-patient error and the Kenward-Roger approximation will be used to estimate the degrees of freedom. If the model assuming AR(1) covariance structure fails to converge, the following covariance structures will be tested one after the other in the order listed until one converges (remaining covariance structures will not be tested): heterogeneous autoregressive order 1 (ARH(1)), compound symmetry, heterogeneous compound symmetry (CSH), Toeplitz (TOEP), and heterogeneous Toeplitz (TOEPH).

7. Output presentations

Appendix 1 shows conventions for presentation of data in outputs.

The templates provided with this SAP describe the presentations for this study and therefore the format and content of the summary tables, figures, and listings to be provided by eXYSTAT.

8. Disposition and withdrawals

The total number of subjects enrolled, the numbers and percentages of randomized subjects (ITT population) who were treated in the study, who completed the 1.5-year visit will be provided with the numbers and percentages of randomized subjects who withdrew prematurely, and a breakdown of the corresponding reasons for withdrawal will be presented by treatment group and overall.

The numbers and percentages of randomized subjects who enrolled in the Follow-Up Period 2, who completed the 3-year visit, the 4-year visit, and Post-Treatment Follow-Up Period 2 (5-year visit) will be provided by treatment group and overall, together with the numbers and percentages of randomized subjects who withdrew from the study prematurely, with a breakdown of the corresponding reasons for withdrawal.

The number and percentage of randomized (ITT population) subjects in the mITT, PP, and safety populations will also be provided by treatment group and overall, together with the numbers and percentages of randomized subjects excluded from the analysis populations, with a breakdown of the corresponding reasons for exclusion.

A listing of subjects not included in the scheduled time windows will be provided for each visit concerned.

All major protocol deviations will be identified prior to unmasking for analysis. A supportive analysis of the primary endpoint will be performed based on the per-protocol analysis population. Subjects with major protocol deviations that can potentially interfere with the interpretation of the treatment effect will not be included in the PP analysis population. A listing showing protocol deviations for each subject will be provided by treatment group.

9. Demographic and other baseline characteristics

Demographic data and other baseline characteristics of the ITT populations will be summarized by treatment group and overall using descriptive statistics.

No statistical tests will be performed to examine any of the demographic or other baseline characteristics for a statistically significant difference between the treatment groups.

The following baseline characteristics will be reported for this study:

- Relevant medical and surgical history
- Prior medications (see section 13 below)
- Duration of disease at baseline (corresponding to the duration of vision loss at baseline for first affected eye): overall and by age classes at consent <15, [15-18[, [18-60[, ≥60-year-old
- Duration of vision loss at baseline of best eye
- Duration of vision loss at baseline of worst eye
- Duration of vision loss of first eye
- Duration of vision loss of second eye
- Time interval between first and second eye vision loss
- Prior use of Idebenone
- Duration of use of Idebenone
- Dose of Idebenone
- Time between stop date of Idebenone and baseline
- Bilateral/Unilateral affected eye status at baseline:

The following demographic characteristics will be reported for this study:

- Age at consent (years)
- Number of patients with age at screening <15, [15-18[, [18-60[, ≥60-year-old
- Age at onset of the disease
- Number of patients with age at onset of the disease <15, [15-18[, [18-60[, ≥60-year-old
- Gender
- Tobacco use
- Alcohol consumption.

Variables derivations:

- Age at consent (years)

Age at Consent = the largest whole number < (Date of informed consent-Date of birth)/365.25.

- Age at onset of the disease (years)

Age at onset of the disease the largest whole number < (Earliest date of onset of vision loss -Date of birth)/365.25.

- Alcohol consumption

For former and current use:
No use
≤1 drinks per day or occasionally
]1-2] drinks per day
>2 drinks per day

- Bilateral/Unilateral affected eye status at baseline

Unilateral eye affected: if LogMAR ≤ 0 for one of the two eyes on the day of injection

Bilateral eyes affected will be derived as: if LogMAR > 0 for both eyes on the day of injection

- Duration of disease at baseline (months)

Duration of disease at baseline =date of baseline – earliest date of onset of vision loss

- Duration of vision loss at baseline for one eye (months)

Duration of vision loss at baseline for one eye = date of baseline – date of onset of vision loss for the considered eye

Demographic and baseline characteristics will be summarized using descriptive statistics. Categorical variables (e.g. gender) will be summarized using frequencies (count and percentages) while quantitative variables (e.g. age, tobacco use) will be summarized using number of subjects with available data (n), mean, SD, median, first quartile (Q1), third quartile (Q3), minimum, and maximum.

These data will be listed by treatment group.

10.Surgical and medical history

Medical and surgical history conditions are defined as those conditions which stop prior to or at Screening and will be coded to the system organ class (SOC) and preferred term (PT) using Medical Dictionary for Regulatory Activities (MedDRA) version 21.1 or later.

Medical and surgical history data of the safety analysis set will be summarized by SOC and PT using counts and percentages based on the number of subjects in each treatment group and overall.

11.Medications

The World Health Organization (WHO) drug dictionary version 01SEP2017EB3 or later will be used for medications coding. The Anatomical Therapeutic Chemical (ATC) classifications ATC2 and ATC4 will be used.

Prior and concomitant medications are defined as follows:

- Prior medications are defined as prescription and non-prescription medications and preparations, including health and/or dietary supplements, prior to first IVT.
- Concomitant medications and therapies are defined as prescription and non-prescription medications and preparations, health and/or dietary supplements other than randomized study treatment that the subject receives during the course of the study (i.e. any time after first IVT). This includes any medication started prior to first IVT and continued after first IVT.

See Appendix 2 for handling of partial dates for medications, in the case where it is not possible to define a medication as prior or concomitant, the medication will be classified by the worst case, i.e. concomitant.

Concomitant treatments will be summarized by coded terms using number and percentage of Safety population subjects by treatment group reporting the medication at least once during the study.

Incidence of corticosteroid therapy will also be presented by route of administration (local corticosteroid alone / Systemic corticosteroid alone / Local and systemic corticosteroid) for the first 6 months of follow-up and for the follow-up between 6 months and 1 year and for each year of follow-up after the first year. A patient will be counted in the "local and systemic corticosteroid" class if he took, during the time windows of interest, at least one local corticosteroid and at least one systemic corticosteroid, whatever the concomitance of these two treatments.

Description of local corticosteroid therapy will be presented using descriptive statistics:

- Duration of treatment (days),
- Reason of use of local corticosteroid (Preventive of ocular inflammation / for AE)
- Indication of use (Ocular/non-ocular)

Description of systemic corticosteroid therapy will be presented using descriptive statistics:

- Duration of treatment (days),
- Daily dose, for systemic corticosteroid therapy (expressed in dexamethasone dose equivalent)

INP	Equivalent Daily Doses	Factor to apply to convert in dexamethasone dose equivalent
Dexamethasone	0.75 mg	1
Cortisone	25 mg	0.75/25
Hydrocortisone	20 mg	0.75/20
Prednisone/Prednisolone	5 mg	0.75/5
Methylprednisolone	4 mg	0.75/4
Triamcinolone	4 mg	0.75/4
Betamethasone	0.75 mg	1
Cortivazol	0.3 mg	0.75/0.3

- Reason of use of systemic corticosteroid (Preventive of ocular inflammation / for AE)

Description of idebenone will be presented using descriptive statistics:

- Duration of treatment (days),
- Daily dose

Description of intraocular pressure lowering agents will be presented using descriptive statistics:

- Duration of treatment (days),

- Reason of use of intraocular pressure lowering agents (Preventive of intraocular pressure increase / for AE)

Listings of patients on corticosteroid therapy, Idebenone and intraocular pressure lowering agent will be provided.

12.Study medication exposure

In this study, each eye will undergo administration of the allocated treatment (GS010 or Placebo) with a single intravitreal injection, performed as a separate procedure. Therefore, no summary table will be produced for the exposure data. eCRF data related to GS010 administration will be listed.

13.Efficacy outcomes

All primary efficacy analyses will be conducted using the ITT population and on the primary eye, which is the second-affected/not-yet-affected eye. Data will be summarized at each visit and presented in tables; change from baseline will be calculated and the associated two-sided 95% confidence interval presented.

13.1. Primary efficacy

The primary efficacy endpoint is the change from baseline of the Best Corrected Visual Acuity (BCVA) reported with LogMAR at 1.5 years post-treatment, in second-affected/not-yet-affected eyes of *ND4* LHON subjects with vision loss up to one year. The change from baseline (Visit 2) in primary (second-affected/not-yet- affected) eyes receiving GS010 and placebo will be the primary response of interest. LogMAR BCVA will be used to represent BCVA.

The definition of the primary eye (the second/not yet affected eye) is based on the vision loss duration. The second/not yet affected eye is the eye with the smallest vision loss duration or with a vision loss duration equal to 0. For subjects who report simultaneous onset of vision loss in both eyes, the right or left eye will be selected randomly to serve as the second-affected/not-yet-affected eye. If the second affected eye did not receive the correct treatment (as set in the randomization scheme) then the subject will be assigned to the arm corresponding to the randomized treatment.

13.1.1. Primary efficacy variable & derivation

Visual acuity (EDTRS and LogMAR)

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

One Early-Treatment Diabetic Retinopathy Study (ETDRS) line equates to 0.1 LogMAR, and 1 ETDRS letter equates to 0.02 LogMAR.

On chart and off chart

On-chart subjects are able to read at least three ETDRS letters on a single line of the ETDRS chart at 1 meter distance and will thus have an ETDRS score, Snellen visual acuity, and LogMAR visual acuity based on the letters/lines read on the ETDRS chart at 4- and 1-meter distances.

Off-chart subjects are not able to read at least three ETDRS letters on a single line of the ETDRS chart at 1 meter distance and thus will not have an ETDRS score. Off-chart subjects will be tested for count fingers (CF) or hand motion (HM) visual acuity per the SOP for Ocular and Vision Testing. The LogMAR visual acuity will be assigned to the validated Lange scale LogMAR value for CF and HM (C. Lange et al. Graefes Arch Clin Exp Ophthalmol (2009) 247:137-142):

- +2.0 LogMAR for CF
- +2.3 LogMAR for HM

For light perception (LP) +4.0 LogMAR is assigned

For no light perception (NLP) +4.5 LogMAR is assigned.

13.1.2. Timing of primary analysis

Primary endpoint analysis will be performed as soon as the last subject completes the 1.5-year LogMAR assessment and data are cleaned. All subjects will be recruited and followed for at least 1.5 years except in case of premature withdrawal. The study will be formally unmasked for the efficacy analysis at 1.5 years to allow an early registration, if results meet the expectations. The study will continue to its scheduled end (5 Years).

13.1.3. Handling of missing data

To address the impact of COVID-19 on study integrity, the LogMAR of Patients without assessment in the protocol-defined window of time [518-699 days] will be imputed by a LogMAR defined by clinical decision during the data-review before the database lock. If data remains missing, it will be imputed with Last observation carried forward method (LOCF).

Number of missing data and a listing with imputed values will be provided.

13.1.4. Primary analysis of primary efficacy variable(s)

The primary analysis will be performed on the ITT population on the planned treatment (as randomized). The primary endpoint is the change from baseline of the BCVA, reported in LogMAR at 1.5 years post-treatment. Descriptive statistics will be presented for the baseline and 1.5 years post-treatment timepoints as well as for the change from baseline to the 1.5 years post-treatment based on data after imputation.

The treatment difference will be assessed with an ANCOVA model which includes terms to adjust for the baseline LogMAR value and includes treatment group as a fixed effect.

The analysis will be performed using a significance level of 0.05 (two-sided).

SAS code:

ARM =subject Treatment arm; change=Logmar change from baseline of second-affected/not-yet-affected eyes; Logmar0=Logmar at baseline

```
proc mixed data=XXX_;
class ARM;
model change=ARM logmar0 /s CL;
estimate " GS010 arm vs Placebo arm" ARM -1 1;
run;
```

13.1.5. Supportive analysis of primary efficacy variable(s)

Shift table and Mac Nemar test between Eyes on-chart/off-chart for LogMAR BCVA at baseline and Eyes on-chart/off-chart for LogMAR BCVA at 1.5 years by arm will be performed.

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

1. Using the mITT population, if different from the ITT population
2. Using the PP population
3. Same analyses will be conducted on the LogMAR instead of the change of LogMAR (analyses will be not adjusted on LogMAR at baseline but with duration from vision loss)

Using a model at the subject level, with repeated values for eye. We need to have comparison between treatment group (GS010 vs Placebo) and comparison between GS010 eye in GS010/ GS010 ARM and GS010 eye in GS010/ Placebo ARM.SAS code:

ID=identifiant; TREAT =(Placebo in GS010/ Placebo ARM / GS010 in GS010/ Placebo ARM / GS010 first affected eye in GS010/ GS010 ARM / GS010 second affected eye in GS010/ GS010 ARM ; change=Logmar change from baseline; Logmar0=Logmar at baseline;

```
proc mixed data=XXX ;
class ID TREAT(ref="PLACEBO") ;
model change= TREAT logmar0 /s CL;
random intercept/sub=ID type=AR(1);
lsmestimate TREAT " GS010 vs Placebo " 1 1 1 -3/cl divisor=3;
lsmestimate TREAT " GS010 in GS010/GS010 ARM vs GS010 in GS010/Placebo ARM" -2
1 1 0/cl divisor=2;
run;
```

4. Primary ITT analysis adjusted on potential confusion factors:

ANCOVA will be used to analyze the effect of treatment on the change from baseline of the BCVA, reported in LogMAR at 1.5 years post-treatment adjusted on baseline LogMAR value, and potential confusion factors. A backward selection approach will be considered to keep significant covariates among: Gender, Age at onset of the disease, Duration of vision loss, Pediatrics/Adults, Previous Idebenone treatment, BCVA at baseline as a continuous parameter, BCVA on-chart/off-chart at baseline, BCVA on-chart/off-chart at nadir, LogCS at baseline as a continuous parameter, LogCS on-chart/off-chart at baseline, LogCS on-chart/off-chart at nadir, baseline GCL volume, baseline Quadrant Temporal RNFL, baseline PMB RNFL and Region. We will consider a p-value of <0.2 in univariate model (adjusted on logMAR at baseline) for entry in the multivariate model and a p-value of >0.05 for removal from multivariate model. If a covariable has a p-value>0.05 but

causes a strong variation in the estimates of the treatment (+/-20%), then it will be conserved in the model and considered as a confounding factor. If linear hypothesis between change from baseline and quantitative variable is not verified, thus quantitative variable should be included in the model as qualitative variable (on the covariate's quartiles for example). Multicollinearity of independent variables (those which enter in the multivariate model) was checked via the variance inflation factor (VIF) statistic. If VIF>10 between two or several covariables, only one covariable should be included in the model.

SAS code:

ARM =subject Treatment arm; change=Logmar change from baseline of second-affected/not-yet-affected eyes; Logmar0=Logmar at baseline; VARi=quantitative covariables i; VARCi=qualitative covariables i

Univariate analysis for quantitative variable:

```
proc mixed data=XXX;
model change= logmar0 VARi/s CL;
run;
```

Univariate analysis for qualitative variable:

```
proc mixed data=XXX;
class VARCi;
model change= logmar0 VARCi /s CL;
lsmeans  VARCi;
run;
```

Multivariate analysis:

```
proc mixed data=XXX;
class ARM VARC1 VARC2 …;
model change=ARM logmar0 VARC1 VARC2… VAR1 VAR2 …/s CL;
run;
```

Test of multicollinearity:

```
proc reg data= XXX;
model change= ARM logmar0 COUNTRY VAR1 VAR2 VAR3 /VIF TOL COLLIN;
run;
```

13.1.6. Sensitivity analysis of primary efficacy variable(s)

Sensitivity analyses of the primary endpoint will be conducted only on observed data (data without imputed values) up to 1.5 years using a MMRM model with baseline adjustment; treatment group, study time points, and treatment group-by-time point interaction as fixed effects; and subjects as random effect.

SAS code:

ID=identifiant; ARM =subject Treatment arm (1=GS010 – 0=Placebo); change=change from baseline of second-affected/not-yet-affected eyes; Logmar0=Logmar at baseline; time=timepoint (time=0 is baseline)

```
proc mixed data=XXX;
where time>0;
class ID time ARM;
model change= time ARM ARM*time logmar0 /CL ddfm=KENWARDROGER;
```

```
repeated time/sub=ID type=AR(1);
lsmeans ARM*time/diff;
run;
```

13.2. Secondary efficacy

Secondary efficacy endpoints will be on ITT population.

13.2.1. Secondary efficacy variables & derivations

Change in LogMAR BCVA

The first secondary efficacy endpoint is the change from baseline in LogMAR BCVA to each follow-up visit (after 1.5 years) (see section 13.1.4) with the same missing data imputation strategy.

For 5-year analysis, additional supportive analyses will be done:

- Using last available LogMAR value

- Same analysis on PP5Y population.

Response status

The second secondary efficacy endpoint is the responder rate.

For subjects discontinued prior to the time of analysis or with missing data, LOCF imputed data will be used.

Three definitions of responders are given below:

***Definition 1 – Gainer eyes**:*

Eyes whose LogMAR BCVA improves (i.e. decrease) by ≥ 0.3 LogMAR (equivalent to a gain of ≥ 15 ETDRS letters) compared to baseline, i.e. an eye is a responder if: LogMAR at interest timepoint – LogMAR at Baseline ≤ -0.30.

***Definition 2 – Stabilized eyes**:*

Eyes whose LogMAR BCVA does not worsen (i.e. does not increase) by ≥ 0.3 LogMAR (equivalent to an eye that loses ≤ 15 ETDRS letters) compared to baseline, i.e. an eye is a responder if: LogMAR at interest timepoint - LogMAR at Baseline ≤ 0.30.

***Definition 3 - Not legally blind**:*

Eyes whose LogMAR visual acuity is < 1.0 at interest timepoint (i.e. better than LogMAR 1.0, equivalent to better than Snellen acuity of 20/200).

This third definition of interest is exploratory given the precipitous vision loss experienced by the majority of *ND4* LHON subjects. Any significant difference in the percentage of second-affected/not-yet-affected responder eyes, receiving GS010 compared to placebo, is pertinent considering the seriousness of the visual impairment.

The response status will be assessed at eye and subject level.

For subject response, the best response observed in either eye will be considered (LogMAR measured on both eyes will not be considered).

Parameters measured with SD-OCT

The changes from baseline at each time point in the following parameters as measured by the SD-OCT:

- Total and Temporal Quadrant Retinal Nerve Fiber Layer (RNFL) thickness;
- Papillo Macular Bundle (PMB) RNFL thickness
- Ganglion Cell Layer macular volume
- ETDRS macular volume

HVF 30-2

The changes from baseline at each time point the following parameters of standardized, automated visual fields obtained with the HVF Analyzer II (HVF 30-2):

- Mean deviation
- Foveal threshold sensitivity
- Pattern standard deviation

If HVF is off then set to missing, if <0 set to 0.

The changes over time of each data point of HVF, with the following thresholds of gain or loss: ≥7 dB; 5-6.99 dB; 3-4.99 dB; ≤2.99 dB. Graphic mapping of HVFs per eye and per patient will be provided. Only reliable HVFs will be included in this analysis.

Contrast sensitivity measured with the Pelli-Robson Low Vision Contrast Sensitivity.

The last set of the 3rd secondary efficacy endpoint is the change from baseline at each time point in low vision contrast sensitivity as measured by the Pelli-Robson chart.

On-chart and off-chart contrast sensitivity

The Contrast Sensitivity score is calculated based on the last triplet of letters for which at least two of the three letters were read correctly (any result obtained past the triplet where the patient could not read at least 2 letters correctly is not to be recorded as the result). On-chart subjects are able to read at least two letters of the first triplet of the Pelli-Robson chart, and will thus have a CS value expressed in LogCS. Off-chart subjects are not able to read at least two letters of the first triplet of the Pelli-Robson chart, and will thus have no CS value. As no equivalent score exists for off-chart LogCS, 0 be assigned for eyes with off-chart LogCS.

Same analyses will be conducted on the LogCS instead of the change of LogCS for all eyes.

Quality of Life assessments

Visual Functioning Questionaire-25

The VFQ-25 Questionnaire is a 25-item version of the 51-item National Eye Institute Visual Function Questionnaire. The VFQ-25 consists of a base set of 25 vision-targeted questions representing 11 vision- related constructs, plus an additional single-item general health rating question. All items are scored from 0 to 5 so that a higher number represents better functioning. To score the questionnaire, each item is converted to a 0 100 scale. Items within each subscale are averaged to create 12 subscale scores. The subscale scores (excluding the general health rating question) are then averaged to calculate the composite score (Mangione 2001), sub- scales with at least one item answered can be used to generate a sub-scale score. Hence, scores represent the average for all items in the sub-scale that the respondent answered. Items that are left blank (missing data) are not taken into account when calculating the scale score.

By averaging the subscale scores rather than the individual items it will give equal weight to each subscale, whereas averaging the items would give more weight to scales with more items.

The secondary efficacy endpoint for VFQ-25 will be the change from baseline in the VFQ-25 composite score to each post-treatment follow-up visit.

Details of the conversion scale, sub-scales, calculations of the subscale scores and formula for averaging the subscale scores can be found in Appendix 4.

36- Item Short Form Health Survey, Version 2

The SF-36 Questionnaire is a generic, subject-reported outcome instrument used to assess QoL. The SF-36 is a 36-question instrument, which assesses 8 health concepts:

1. limitations in physical activities because of health problems;
2. limitations in social activities because of physical or emotional problems;
3. limitations in usual role activities because of physical health problems;
4. bodily pain;
5. general mental health (psychological distress and well-being);
6. limitations in usual role activities because of emotional problems;
7. vitality (energy and fatigue);
8. general health perceptions.

The scale score of each domain is calculated based on the summed score across items included in the domain and is rescaled to 0 to 100 with higher scores indicating better health states. The rescaled score of each domain is then normalized for the US population. Finally, the Physical Component Score (PCS) and Mental Component Score (MCS) will be computed using the normalized domain scores (Ware 2007).

Items that are left blank (missing data) are not taken into account when calculating the scale scores. Hence, scale scores represent the average for all items in the scale that the respondent answered.

The secondary efficacy endpoints for the SF-36 will be the change from baseline in each of the SF-36 domains as well as in the two composite scores to each post-treatment follow-up visit.

Details of the conversion rules, transformation scales, scale raw score range and scoring rules for overall component can be found in Appendix 5.

13.2.2. Missing Data Methods for Secondary Efficacy Variable(s)

Missing data will be imputed using the control-based imputation described in section 13.1.3 for BCVA measured in LogMAR,

Total score and individual subscale scores form the VFQ-25, composite score and individual domain score as well as categorical secondary efficacy endpoints will not be imputed.

13.2.3. Analysis of Secondary Efficacy Variables

Change in LogMAR BCVA analysis

Change in LogMAR compared to baseline, at each timepoint of the Follow-Up Period up to 5 years post- treatment will be analyzed at each visit using separate ANCOVA models as primary endpoint (see section 13.1.4).

Response status

For each response parameter and at each timepoint of the Follow-Up Period up to 5 years post- treatment, logistic model, including a factor for treatment group and the baseline value as a covariate will be considered on:

- Second-affected eye at timepoint (at eye level)
- First-affected eye at timepoint (at eye level)
- First and second affected eye at timepoint (at subject level): best response between both eyes

SAS code:

```
ARM =subject Treatment arm (1=GS010 – 0=Placebo); resp=responder variable; Logmar0=Logmar at baseline;
proc logistic data=XXX;
class ARM;
model resp=logmar0 ARM;
run;
```

For each of the response definitions and at each timepoint of the Follow-Up Period up to 5 years post- treatment, a logistic model, including a factor for treatment group, eye, the baseline value as a covariate, will be fit using GLMM for first and second affected eye at timepoint (at eye level)

SAS code:

```
ARM =subject Treatment arm (1=GS010 – 0=Placebo); resp=responder variable; Logmar0=Logmar at baseline;
proc glimmix data=XXX_ana;
class ID ARM;
model resp= ARM logmar0 /dist=bin solution cl;
random intercept/sub=ID type=ar(1);
run;
```

Parameters measured with SD-OCT, HVF and contrast sensitivity

The parameters detailed in sections 13.2.1 will be analyzed separately using a MMRM model, with adjustment for baseline score, as described in section 13.1.6.

Quality of Life assessments

Visual Functioning Questionnaire-25 (VFQ-25)

VFQ25 measures will be assessed at the subject level, looking at change from baseline for all subjects and comparing the change from baseline between the two arms (bilaterally treated subjects vs. unilaterally treated subjects).

Composite score and individual subscale scores from VFQ-25 will be summarized using descriptive statistics. Change from baseline at each visit for the composite score and individual subscales scores will be analyzed separately using a MMRM model, with adjustment for baseline score, as described in section 13.1.6.

36-Item Short Form Health Survey, Version 2 (SF-36v2)

SF-36 measures will be assessed at the subject level, looking at change from baseline for all subjects and comparing the change from baseline between the two arms (bilaterally treated subjects vs. unilaterally treated subjects).

Composite scores and individual domains from SF-36v2 will be summarized using descriptive statistics. Change from baseline at each visit for the composite scores and individual domains will be analyzed separately using a MMRM model, with adjustment for baseline score, as described in section 13.1.6.

13.3. Other efficacy

Nadir for LogMAR BCVA and for LogCS

Nadir of BCVA/LogCS is defined for each eye of each subject as the worst value observed from baseline to the post-treatment timepoint of interest, including baseline and the post-treatment timepoint values.

For LogCS only, if there is at least one value off-chart between baseline and timepoint, then Nadir cannot be calculated as no off-chart value exists; off-chart LogCS will be imputed a value of 0.

If there are more than one off-chart value from baseline to the timepoint of interest, then the first value over time will be considered the nadir. If there is more than one worst value from baseline to the point of interest (i.e. identical worst values occurring more than once), then the first value over time will be considered the nadir.

The change from Nadir BCVA/LogCS will be calculated for each eye of each subject and will be analyzed as primary endpoint (see section 13.1.4).

The time to Nadir BCVA/LogCS will be calculated for each eye of each subject and will be expressed in Weeks. This analysis will be only descriptive.

Clinically Relevant Recovery (CRR)

At the eye level, a CRR from baseline is defined as:

- For eyes on-chart at baseline, an improvement of at least -0.2 LogMAR
- For eyes off-chart at baseline, eyes which became on-chart.

On-chart value is defined as ≤ LogMAR 1.6

Subject with CRR from baseline is defined as subject with a CRR in at least one eye: Subject with at least one eye which was on chart at baseline, and which had an improvement of at least -0.2 LogMAR from baseline, or which was off-chart at baseline but became on-chart.

Subjects with CRR from baseline will be assessed for all subjects, bilaterally GS010 treated subjects and unilaterally GS010 treated subjects.

At the eye level, a CRR from nadir is defined as:

- For eyes on-chart at nadir, an improvement of at least -0.2 LogMAR from nadir
- For eyes off-chart at nadir, eyes which became on-chart.

Subject with CRR from nadir is defined as subject with a CRR in at least one eye: Subject with at least one eye which was on chart at nadir, and which had an improvement of at least -0.2 LogMAR from nadir, or which was off-chart at nadir but became on-chart.

Subjects with CRR from nadir will be assessed for all subjects, bilaterally GS010 treated subjects and unilaterally GS010 treated subjects.

These responder endpoints will be analyzed as section 13.2.3

Other responses parameters

Following responder endpoints will be analyzed as section 13.2.3:

- Eyes/subjects whose LogMAR BCVA improves (i.e. decrease) by ≥ 0.2 LogMAR (equivalent to a gain of ≥ 10 ETDRS letters) compared to baseline, i.e. an eye/subject is responder if:

  LogMAR at the timepoint of interest - LogMAR at Baseline ≤ -0.20
- Eyes/subjects whose LogMAR BCVA improves (i.e. decrease) by ≥ 0.3 LogMAR (equivalent to a gain of ≥ 15 ETDRS letters) compared to baseline, i.e. an eye/subject is responder if:

  LogMAR at the timepoint of interest - LogMAR at Baseline ≤ -0.30
- Eyes/subjects whose LogMAR BCVA improves (i.e. decrease) by ≥ 0.3 LogMAR (equivalent to a gain of ≥ 15 ETDRS letters) compared to nadir, i.e. an eye/subject is responder if:

  LogMAR at the timepoint of interest - LogMAR at Nadir ≤ -0.30
- Eyes/subjects whose LogMAR visual acuity is ≤ 1.0 (i.e. equal or better than LogMAR 1.0, equivalent to equal or better than Snellen acuity of 20/200).
- Eyes/subjects whose LogMAR visual acuity is < 1.0 (i.e. better than LogMAR 1.0, equivalent to better than Snellen acuity of 20/200).
- Eyes/subjects whose LogMAR visual acuity is ≤ 1.6 (i.e. equal or better than LogMAR 1.6, equivalent to equal or better than Snellen acuity of 20/800).

- Eyes/subjects whose LogCS improves by ≥ 0.2 LogCS compared to baseline, i.e. an eye/subject is responder if:

  LogCS at the timepoint of interest - LogCS at Baseline ≥ 0.20.

- Eyes/subjects whose LogCS improves by ≥ 0.3 LogCS compared to baseline, i.e. an eye/subject is responder if:

  LogCS at the timepoint of interest - LogCS at Baseline ≥ 0.30.

- Eyes/subjects whose LogCS improves by ≥ 0.3 LogCS compared to nadir LogCS, i.e. an eye is responder if:

  LogCS at the timepoint of interest - LogCS at Nadir ≥ 0.30.

On-chart/off-chart BCVA and LogCS

Shift table and Mac Nemar test between Eyes on-chart/off-chart for BCVA/LogCS at baseline and Eyes on-chart/off-chart for LogMAR BCVA/LogCS at each timepoint (1.5, 2, 3, 4, 5 years) in each eye group will be performed.

Shift tables: Off-chart BCVA eyes at baseline moving to on-chart BCVA; On-chart BCVA eyes at baseline stay on-chart BCVA; Off-chart LogCS eyes at baseline moving to on-chart LogCS; On-chart LogCS at baseline stay on-chart LogCS

Exploratory efficacy

- Primary and secondary analyses will be repeated for the first affected eyes as exploratory efficacy analyses.

14.Safety outcomes

All outputs for safety outcomes will be based on the safety population. Period up to 1.5 year (0-1.5 year) and up to 5 years (0-5 years) will be presented separately.

Adverse events will be summarized as outlined above, permitting an assessment of change over the 5 years of follow-up. Other safety endpoints will be summarized as outlined above, showing change from baseline, for continuous endpoints.

There will be no statistical comparisons of the two treatment groups with respect to safety, unless otherwise specified within the relevant section.

If the second affected eye did not receive the correct treatment (as set in the randomization scheme) then the subject will be assigned to the arm corresponding to the treatment actually received. In case of assignment error, the treatment prevails over the order of affection of eye (i.e. first or second-affected/not- yet-affected eye corresponding to secondary or primary eye, respectively).

14.1. Adverse events

An adverse event (AE) is defined as any untoward medical occurrence in a subject or clinical investigation subject administered a pharmaceutical product and which does not necessarily have a causal relationship with this

treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether related to the medicinal (investigational) product, or not. Any event or laboratory abnormality that leads to a medical intervention, including withdrawal of IMP or significant additional concomitant therapy, will be considered an AE. Worsening of visual acuity determined by the Investigator to be due to progression of LHON will not be considered an AE.

AEs should be volunteered by the subject, be observed from examination of the subject at a clinic visit or be from observations of clinically significant laboratory values or other/special examination abnormal values. AEs will not be solicited by use of a specific list of anticipated events.

AEs will be coded to the SOC and PT using MedDRA central coding dictionary, Version 21.1 or higher.

Treatment-emergent adverse events (TEAEs) are defined as AEs that started on or after study treatment administration at Visit 3 or that represent an exacerbation of a condition that is present at baseline after treatment administration. See Appendix 2 for handling of partial dates for AEs. In the case where it is not possible to define an AE as treatment emergent or not, the AE will be classified by the worst case; i.e. treatment emergent.

Local (i.e. ocular) and systemic (i.e. non-ocular) AEs will be summarized separately. Ocular TEAEs will be summarized by treatment group for the first affected eye and second-affected/not-yet-affected eye separately. Systemic TEAE data will be summarized for the 2 treatment groups combined.

The following summaries will be produced for systemic AEs and ocular AEs separately using counts and percentages, i.e. number of subjects and percentage of subjects.

- Overall summary of AEs for the entire study, including the number of events reported
- TEAEs, including the number of events reported
- TEAEs by relationship to study treatment (‘Unrelated’, ‘Unlikely’, ‘Related’)
- TEAEs by relationship to study procedure (‘Unrelated’, ‘Unlikely’, ‘Related’)
- TEAEs by severity
- TEAEs leading to study discontinuation
- Serious TEAEs
- TEAEs leading to death, life threatening, hospitalization and other
- TEAEs of Special Interest (AESI)
- TEAEs of Special Interest (AESI) with IOP
- TEAEs of Special Interest (AESI) with intra inflammation

The following summaries will be produced for systemic AEs and ocular AEs separately using counts and percentages, i.e. number of subjects and percentage of subjects.

- TEAEs
- TEAEs related to study medication
- TEAEs related to study procedure

- Severe TEAEs
- TEAEs leading to study discontinuation
- Serious TEAEs
- TEAEs leading to death, life threatening, hospitalization and other
- TEAEs of Special Interest (AESI)
- TEAEs of Special Interest (AESI) with IOP
- TEAEs of Special Interest (AESI) with intra inflammation

Subjects are only counted once within each SOC or PT. Subjects with multiple events within a particular SOC or PT will be counted under the category of their closest drug-related event or procedure-related event within that SOC or PT for the summary tables by closest relationship to study drug and study procedure, respectively.

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

Within each systemic AE frequency table, the SOCs will be ordered in alphabetical order, and the PTs will be listed in descending order of frequency.

Within each ocular AE frequency table, the SOCs will be ordered in alphabetical order, and the PTs will be listed in descending order of frequency in the second-affected/not-yet-affected eye of the GS010 treatment group.

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

A summary of SAEs will be provided, grouped by system organ class and preferred term.

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

In addition, a summary of AEs of intraocular inflammation and the concomitant medications administered for such events will be summarized, including the time course of the events and the concomitant medications.

All AEs recorded on the eCRF will be listed.

14.1.1. All TEAEs

Incidence of TEAEs related to study medication, TEAEs related to study procedure, and severe TEAEs will be presented by SOC and PT. A further tabulation presenting the PTs in descending order of frequency will also be presented.

Severity

Severity will be classified as mild/ moderate/ severe (increasing severity). TEAEs starting after the first dose of study medication with a missing severity will be classified as severe.

Relationship to Study Medication and to Study Procedure

Relationship, as indicated by the Investigator, will be classified as '*Unrelated'*, '*Unlikely'*, '*Possible'* or '*Probable'* (increasing severity of relationship), where:

- *Unrelated*: There is evidence of relationship to a cause other than the IMP/study procedures. Does not meet criteria listed under unlikely, possible or probable.
- *Unlikely*: Does not follow a reasonable temporal sequence from administration. Is most likely produced by the subject's clinical state or by environmental factors or other therapies administered.
- *Possible*: Follows a reasonable temporal sequence from administration. Is not likely produced by the subject's clinical state or by environmental factors or other therapies administered.
- *Probable*: Follows a reasonable temporal sequence from administration. Clear-cut temporal association with IMP.

A "related" TEAE is defined as a TEAE with a relationship to study medication and/or study procedure of '*Possible'* or '*Probable'*. TEAEs starting after the first dose of study medication with a missing relationship to study medication and/or study procedure will be assigned a relationship of *'Probable'*.

14.1.2. Adverse events leading to discontinuation from study

TEAEs leading to discontinuation from study are those events recorded as "Withdrawal from study" on the AE page of the eCRF.

A summary of TEAEs leading to discontinuation from study by SOC and PT will be prepared. A listing of all AEs leading to discontinuation from study (non-TEAEs and TEAEs) will be provided.

14.1.3. Serious adverse events

Serious adverse events (SAEs) are those events recorded as "Serious" on the AE page of the eCRF.

A summary of serious TEAEs by SOC and PT will be prepared. A listing of all SAEs (non-TEAEs and TEAEs) will be provided.

14.1.4. Adverse events leading to death

TEAEs leading to death are those events which are recorded as "Fatal" on the Adverse Events page of the eCRF.

A summary of TEAEs leading to death by SOC and PT will be prepared. A listing of all AEs leasing to death (non-TEAEs and TEAEs) will be provided.

14.1.5. Adverse events of special interest

The following AEs have been identified as AESI:

- Intraocular inflammation AE defined as any TEAE of the following MedDRA terms (all terms are PTs unless noted as an LLT):
  - Anterior chamber inflammation (anterior chamber cell, anterior chamber fibrin, anterior chamber flare, anterior chamber inflammation, anterior uveitis, aqueous fibrin, cyclitis, hypopyon, iridocyclitis, iritis, keratic precipitates)
  - Intermediate inflammation (autoimmune uveitis, pars planitis (LLT), vitreal cells, vitritis, vitreous haze, intermediate uveitis, intermediate inflammation, intermediate uveitis (LLT))
  - Posterior inflammation (choroiditis, chorioretinitis, retinitis, posterior inflammation)

- Non-specified eye inflammation (eye inflammation, uveitis)
- IOP increase

Note that the above terms are based on MedDRA version 21.1 and will be reevaluated and updated based on the most recent MedDRA version available.

In addition, a summary of AEs of intraocular inflammation and the concomitant medications administered for such events will be summarized, including the time course of the events and the concomitant medications.

For intraocular inflammation AE, duration of inflammation and maximal grade of inflammation will be described.

Likewise, a summary of AEs of IOP increase and the concomitant medications administered for such events will be summarized, including the time course of the events and the concomitant medications.

For IOP increase AE, duration of IOP increase will be described.

A patient listing of all ocular AEs will be provided.

14.1.6. Treatment for AEs related to elevation of IOP

Treatment of AEs with elevation of IOP will be described with number of patients using IOP lowering treatment (None/IOP agent lowering local alone/IOP agent lowering systemic alone/ local and systemic in association) for the first 6 months of follow-up and for the follow-up between 6 months and 1 year and for each year of follow-up after the first year.

Duration of treatment will be described.

14.1.7. Treatment for AEs related to intraocular inflammation

Treatment of AEs related to an intraocular inflammation will be described with number of patients using corticoids (None/local CS alone/systemic CS alone/ local and systemic in association) for the first 6 months of follow-up and for the follow-up between 6 months and 1 year and for each year of follow-up after the first year. Number and percent of inflamed eyes with at least 7 days of CS treatment will be described for each type of CS (local CS alone/systemic CS alone/ local and systemic in association) for the first 6 months of follow-up and for the follow-up between 6 months and 1 year and for each year of follow-up after the first year.

14.2. Deaths

If any subjects died during the study, as recorded on the "deaths" page of the eCRF, the information will be presented in a summary table and a by subject data listing.

14.3. Laboratory evaluations

The laboratory assessments that will be included in the reporting are:

- Hematology/serum chemistry
- Liver function tests
- Humoral immune response to AAV2
- Cellular immune response to AAV2

- Biodissemination

Quantitative laboratory measurements reported as "< X", i.e. below the lower limit of quantification (BLQ), or "> X", i.e. above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "< X" or "> X" in the listings.

If the values are equal to "BLQ", "BLD" or "<10" then impute them to 0.

Parameters will be assessed using descriptive statistics.

Hematology tests: Complete blood count including red blood cells, hemoglobin, hematocrit, white blood cells with differential and platelets.

Serum chemistry tests: Glucose, lipase, amylase, calcium, sodium, potassium, chloride, bicarbonate, blood urea nitrogen (BUN), creatinine and albumin.

Liver function tests: aspartate aminotransferase (AST), alanine aminotransferase (ALT), ALP, Gamma Glutamyl transferase (GGT) and total bilirubin.

For these three classes of laboratory tests, number of subjects with clinically significant abnormal values at each time point for each laboratory test will be described. A By-Subject Listing of all values + change for the parameter will be done for patients with at least one clinically significant abnormal value for the parameter. Number of subjects with available data (n), mean, SD, median, Q1, Q3, minimum and maximum at each time point for the absolute laboratory value by treatment groups will also be described.

For Liver function tests only, an additional analysis will be done: for each parameter, the number of patients above 2x(upper limit of normal range of the parameter) and the number of patients above 3x(upper limit of normal range of the parameter) will be described by timepoints and by groups.

Humoral and Cellular Immune Responses Against Recombinant Adeno-Associated Virus 2 (rAAV2) and biodissemination

For each visit: Number patients with positive and negative values will be performed. Positivity will be defined with laboratory.

Only for positive rAAV2 (>0 for Cellular Immune Responses -don't included '<10'- and ≥0 for Humoral Immune Responses): the number of subjects with available data (n), mean, SD, median, Q1, Q3, minimum and maximum at each time point for the absolute laboratory value by groups will be described and a graph of mean value overtime will be provided on a logarithmic scale.

Quantitative laboratory measurements will be compared with the relevant laboratory reference ranges in SI units and categorized as:

- Low: Below the lower limit of the laboratory reference range.
- Normal: Within the laboratory reference range (upper and lower limit included).
- High: Above the upper limit of the laboratory reference range.

14.4. ECG evaluations

The results will be determined to be normal, abnormal but not clinically significant or abnormal clinically significant by the Investigator for screening and for comparison to post-treatment assessments.

The number and percentage of subjects with normal, abnormal but not clinically significant and abnormal clinically significant ECG findings will be summarized by visit. Additionally, shift tables showing changes between baseline and EOS (2 years) will also be produced.

All ECG data will be listed by subject.

14.5. Vital signs

The following vital signs measurements will be reported for this study:

- Seated Systolic Blood Pressure (mmHg)
- Seated Diastolic Blood Pressure (mmHg)
- Seated Pulse Rate (bpm)
- Temperature (ºC)

Systolic and diastolic blood pressures, temperature, and pulse rate will be displayed in summary tables including number of subjects with available data (n), mean, SD, median, Q1, Q3, minimum and maximum at each visit. The descriptive statistics for change from baseline will be presented in the same way.

Individual data listings of systolic and diastolic blood pressure, temperature, and pulse rate will be produced for all subjects over time. The data will appear in ascending order by subject ID, parameter, and date of measurement in this listing.

14.6. Physical examination

Physical examination will include general appearance, skin, head, eyes, ears, nose, throat, neck, thyroid, chest/lungs, heart, abdomen, lymph nodes, extremities, height and body weight.

Physical examination results from the visit will be categorized as normal or abnormal and shifts from baseline to the most abnormal post-baseline assessment will be summarized by body system.

14.7. Other safety assessments

14.7.1. Goldmann Applanation Tonometry

Intra ocular pressure (IOP) measurements in mmHg (including post-injection IOP measurements on the day of injection) will be summarized at all timepoints (included 30 min post injection measure) using summary statistics for continuous variables, by treatment groups for the treated eye and comparative eye separately.

For post-injection measurements and for IOP measured during an ocular inflammation AEs, following statistics will be provided by treatment groups for the treated eye and comparative eye separately:

- Distribution of duration of ocular inflammation AEs,
- Number and percentages of elevated IOP. An elevated IOP is defined as a value of IOP > 22 mmHg,
- Distribution of mean IOP during ocular inflammation AEs
- Distribution of maximum IOP during ocular inflammation AEs,

- Duration that IOP> 22 mmHg (days) defined as date of last value > 22 mmHg – date of first value > 22 mmHg, (if only one value time=1 day)

For eyes with at least one elevated IOP, a listing of all values of IOP for these eyes will be provided. Following information will be provided in the listing: study name, subject number, treatment arm, eye concerned, date of injection, date of examination, time between examination and injection and IOP value.

14.7.2. Ocular and Slit Lamp Examination

The result (Normal/Abnormal) before pupil dilation and after dilation will be summarized by anatomical location using counts and percentages, by treatment groups for the treated eye and comparative eye separately and for each timepoint of interest (including baseline examination).

Description of OIS by treatment groups for the treated eye and comparative eye separately, for each visit. Same analyses will be done on each sub-score:

- Anterior chamber cell score
- Anterior chamber flare score
- Vitreous cell score
- Vitreous haze score

For patients with at least one AE related to an intraocular inflammation, mean OIS of all OIS measured during the AE ocular inflammation will be described for the eye concerned by the AE. Same analyses will be done on each of the 4 sub-scores. For the 4 sub-scores, distribution of each category will also be presented. These analyses will be done for anterior intraocular inflammation and for vitreous intraocular inflammation separately.

15.References

Mangione, C, MD. 2000. "NEI VFQ-25 Scoring Algorithm." National Eye Institute Visual Function Questionnaire-25 (NEI-VFQ-25).

Rubin, D. 1987. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons.

Ware JE, Kosinski M, et al. 2007. *User's Manual for the SF-36v2™ Health Survey (2nd Ed).* Lincoln, RI, Quality Metric Incorporated.

Yan Z, et al. 2009. "Precise gene-dose alleles for chemical genetics." *Genetics* 182 (2): 623-6.

APPENDIX 1. Programming conventions for outputs

DATES & TIMES

Depending on data available, dates and times will take the form yyyy-mm-ddThh:mm:ss.

SPELLING FORMAT

The report as well as tables, figures and listings will be prepared in English (US).

ANALYSIS METHODS

Unless otherwise noted, data for quantitative variables will be summarized by descriptive statistics, i.e. the arithmetic mean, standard deviation (SD), median, first quartile (Q1), third quartile (Q3), minimum and maximum.

Data for binary variables (e.g. responder analysis) will be summarized using counts and percentages.

In general, categorical variables will be summarized using the number and percentage of subjects for each characteristic.

Least square (LS) means, 95% confidence interval and P-values will be derived from models including covariates (see primary model) in addition to treatment.

Rounding conventions:

Minimum and maximum: same as observed data

For the mean, median, Q1, and Q3: one decimal more than in the observed data

For the SD and 95% confidence interval: two decimals more than in the observed data For p-values: Four decimals or "<0.0001" if the p-value is smaller than 0.0001.

LISTINGS

All listings will be ordered by the following (unless otherwise indicated in the template):

- randomized treatment group (or treatment received if it's a safety output), first by active dose
- and then placebo
- center-subject ID,
- parameter
- date of measurement/assessment/event/sample collection (where applicable)
- For listings where non-randomized subjects are included, these will appear in a category after the randomized treatment groups labeled 'Not Randomized'.

APPENDIX 2. Partial date conventions

PARTIAL DATES FOR AES AND CMs:

Partial Start Date

1. If start date year/month are the same as year/month of study med start date, and end date is on or after study med start date then impute the start date to be the study med start date.
2. Otherwise set to earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown).

Partial End Date

Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown)

The following rules will be applied to missing or incomplete start dates when determining if an AE is treatment emergent. 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.

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

START DATE	STOP DATE	ACTION
Known	Known	If start date < study med start date, then not TEAE If start date ≥ study med start date, then TEAE
	Partial	If start date < study med start date, then not TEAE If start date ≥ study med start date, then TEAE
	Missing	If start date < study med start date, then not TEAE If start date ≥ study med start date, then TEAE

START DATE	STOP DATE	ACTION
Partial, but known components show that it cannot be on or after study med start date	Known	Not TEAE
	Partial	Not TEAE
	Missing	Not TEAE
Partial, could be on or after study med start	Known	If stop date < study med start date, then not TEAE If stop date ≥ study med start date, then TEAE

	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, then not TEAE If stop date ≥ study med start date, then TEAE
	Missing	Assumed TEAE
Missing	Known	If stop date < study med start date, then not TEAE If stop date ≥ study med start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, then not TEAE If stop date ≥ study med start date, then TEAE
	Missing	Assumed TEAE

ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS:

START DATE	STOP DATE	ACTION
Known	Known	If stop date < study med start date, assign as prior If stop date ≥ study med start date, assign as concomitant
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, assign as prior If stop date ≥ study med start date, assign as concomitant
	Missing	If stop date is missing could never be assumed a prior medication, assign as concomitant
Partial	Known	Impute start date follows: . If start date year/month are the same as year/month of study med start date, and end date is on or after study med start date then impute the start date to be the study med start date. . Otherwise set to earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown) Then: If stop date < study med start date, assign as prior If stop date ≥ study med start date, assign as concomitant

START DATE	STOP DATE	ACTION
	Partial	Impute start date as follows: . If start date year/month are the same as year/month of study med start date, and end date is on or after study med start date then impute the start date to be the study med start date. . Otherwise set to earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown). Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown) Then: If stop date < study med start date, assign as prior If stop date ≥ study med start date, assign as concomitant
	Missing	Impute start date as follows: . If start date year/month are the same as year/month of study med start date, and end date is on or after study med start date then impute the start date to be the study med start date. . Otherwise set to earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown). Then: If stop date is missing could never be assumed a prior medication, assign as concomitant
Missing	Known	If stop date < study med start date, assign as prior If stop date ≥ study med start date, assign as concomitant
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, assign as prior If stop date ≥ study med start date, assign as concomitant
	Missing	Assign as concomitant

APPENDIX 3. Visual functioning questionnaire-25 (VFQ-25)

Version 2000

The National Eye Institute 25-Item Visual Function Questionnaire (VFQ-25) (Mangione 2000)

Version 2000

This final version of the VFQ-25 differs from the previous version in that it includes an extra driving item from the appendix of supplementary questions as part of the base set of items. Also, the revised scoring algorithm excludes the single-item general health rating question from the calculation of the vision-targeted composite score. Because of these 2 changes, the base set of items actually includes 26 questions, however, only 25 are vision- targeted and included in the composite score. Please see the "Frequently Asked Questions" or FAQ section for additional clarifications of these changes.

Background

The National Eye Institute (NEI) sponsored the development of the VFQ-25 with the goal of creating a survey that would measure the dimensions of self-reported vision-targeted health status that are most important for persons who have chronic eye diseases. Because of this goal, the survey measures the influence of visual disability and visual symptoms on generic health domains such as emotional well-being and social functioning, in addition to task-oriented domains related to daily visual functioning. Questions included in the VFQ- 25 represent the content identified during a series of condition-specific focus groups with patients who had age- related cataracts, glaucoma, age- related macular degeneration, diabetic retinopathy, or CMV retinitis.

The VFQ-25 is the product of an item-reduction analysis of the longer field test version of the survey called the 51-item National Eye Institute Vision Function Questionnaire (NEI-VFQ).[2] The longer version contains 51 questions which represent 13 different sub-scales. The NEI- VFQ Field Test Study collected the data needed to examine the reliability and validity of the survey across all of the above-mentioned ocular diseases.

Also, reliability and validity was assessed in a heterogeneous group of patients with low vision from any cause and a group of age-matched persons with normal vision. A published report describes the psychometric properties of the longer field test version of the survey. [3] Additionally, a number of clinical studies have used either the 51 or the 25- item version of the NEI-VFQ across a number of chronic ocular conditions. [4-8] Despite the success of the longer field test version and its continued use, to enhance feasibility a short-form version was planned since the earliest developmental phase.

The VFQ-25 consists of a base set of 25 vision- targeted questions representing 11 vision-related constructs, plus an additional single-item general health rating question. The VFQ-25 also includes an appendix of additional items from the 51-item version that researchers can use to expand the scales up to 39 total items. All items in the VFQ-25 are from the 51-item field test version; no new items were developed for use in the VFQ-25. Unless otherwise specified, the remainder of this document will use the term VFQ-25 to refer to the base set of items.

The VFQ-25 takes approximately 10 minutes on average to administer in the interviewer format. There is also a self-administered version of the survey, however, psychometric testing of the self- administered version has not been done. The VFQ-25 generates the following vision-targeted sub- scales: global vision rating (1), difficulty with near

vision activities (3), difficulty with distance vision activities (3), limitations in social functioning due to vision (2), role limitations due to vision (2), dependency on others due to vision (3), mental health symptoms due to vision (4), driving difficulties (3), limitations with peripheral (1) and color vision (1), and ocular pain (2). Additionally, the VFQ-25 contains the single general health rating question which has been shown to be a robust predictor of future health and mortality in population-based studies. Please see the FAQ section for more information about the general health rating question.

Development of the NEI VFQ-25

The guiding principles for the selection of the short- form items included: 1) low item-level missing data rates; 2) normal distribution of response choices; and 3) retention of items that explained the greatest proportion of variance in the 51-item sub-scales. The items retained in the VFQ-25 and the optional items (provided in the appendix to the survey) are listed on Table 1. A report describing the performance of the VFQ-25 relative to the Field Test version is currently under review. [2] The reliability and validity of the VFQ-25 is similar to that observed for the 51-item version of the survey. On average, each VFQ-25 sub-scale predicts 92% of the variance in the corresponding 51-item sub- scale score.

Optional Items

Appendix 1 consists of additional questions that users may add to a specific sub-scale. Inclusion of these may be helpful if a particular sub-scale represents the primary domain of vision-targeted HRQOL that is felt to be most important for the condition under study. For example, if a user is testing a new treatment for macular degeneration, by adding near vision questions A3, A4, and A5 to VFQ-25 questions 5, 6, and 7, the investigator would have a six-item near vision scale rather than a three-item scale. The addition of these items would enhance the reliability of the near vision sub-scale and is likely to improve the responsiveness of the sub-scale to the intervention over time (Table 6). If items from the appendix are used, the VFQ-25 developers would encourage users to incorporate all optional items for a given sub-scale. This strategy will enhance the comparability of results across studies.

Scoring

Scoring VFQ-25 with or without optional items is a two-step process:

- First, original numeric values from the survey are re-coded following the scoring rules outlined in Table 2. All items are scored so that a high score represents better functioning. Each item is then converted to a 0 to 100 scale so that the lowest and highest possible scores are set at 0 and 100 points, respectively. In this format scores represent the achieved percentage of the total possible score, e.g. a score of 50 represents 50% of the highest possible score.
- In step 2, items within each sub-scale are averaged together to create the 12 sub-scale scores. Table 3 indicates which items contribute to each specific sub-scale. Items that are left blank (missing data) are not taken into account when calculating the scale scores. Sub-scales with at least one item answered can be used to generate a sub-scale score. Hence, scores represent the average for all items in the sub- scale that the respondent answered.

Composite Score Calculation

To calculate an overall composite score for the VFQ-25, simply average the vision-targeted sub- scale scores, excluding the general health rating question. By averaging the sub-scale scores rather than the individual items we have given equal weight to each sub-scale, whereas averaging the items would give more weight to scales with more items.

Table 1. Item Number Translation from the 51-Item Field Test Version to the VFQ 25

S = retained in the VFQ-25, A = retained in the appendix should be used for the VFQ-39,--- = deleted from the VFQ-25 & VFQ-39

Field Test Version Ques.#	Sub-scale	Status	VFQ-25 Ques. #	Field Test Version Ques.#	Sub-scale	Status	VFQ-25 Ques. #
1	general health	S	1	29	social fx	---	---
2	general health	A	A1	30	social fx	A	A9
3	general vision	S	2	31	social fx	S	13
4	expectations	---	---	32	distance vision	A	A8
5	well-being/ distress	S	3	33	distance vision	A	A7
6	well-being/ distress	---	---	34	distance vision	S	14
7	ocular pain	S	19	35	driving (filter item)	S	15
8	expectations	---	---	35a	driving (filter item)	S	15a
9	expectations	---	---	35b	driving (filter item)	S	15b
10	expectations	---	---	35c	driving	S	15c
11	well-being/ distress	S	25	36	driving	---	---
12	ocular pain	S	4	37	driving	S	16
13	well-being/ distress	---	---	38	driving	S	16a *
14	general vision	A	A2	39a	role limitations	S	17
15	near vision	S	5	39b	role limitations	A	A11a
16	near vision	A	A3	39c	well-being/ distress	---	---
17	near vision	S	6	39d	role limitations	---	---
18	near vision	---	---	39e	role limitations	A	A11b
19	near vision	S	7	39f	role limitations	S	18

20	distance vision	S	8	40	well-being/ distress	A	A12
21	distance vision	---	---	41	dependency	S	20
22	distance vision	S	9	42	well-being/ distress	S	21
23	peripheral vision	S	10	43	well-being/ distress	S	22
24	distance vision	A	A6	44	dependency	---	---
25	social fx	S	11	45	dependency	A	A13
26	near vision	A	A4	46	dependency	S	23
27	color vision	S	12	47	dependency	S	24
28	near vision	A	A5				

* VFQ-25 item 16a was listed in previous versions as part of the appendix of supplemental items (#A10).

Table 2. Scoring Key: Recoding of Items

Item numbers	Change in original response category (a)	To recoded value of:
1,3,4, 15(c)	1 2 3 4 5	100 75 50 25 0
2	1 2 3 4 5 6	100 80 60 40 20 0
5,6,7,8,9,10,11,12,13,14,16,16a A3,A4,A5,A6,A7,A8,A9(c)	1 2 3 4 5 6	100 75 50 25 0 *

17,18,19,20,21,22,23,24,25, A11a,A11b,A12,A13	**1** **2** **3** **4** **5**	**0** **25** **50** **75** **100**
A1,A2	**0 to 10**	**0 to 100**

0 to 100[(a)] Precoded response choices as printed in the questionnaire.

[(b)] Item 15c has four-response levels but is expanded to a five-levels using item 15b. Note: If 15b=1, then 15c should be recoded to "0"

If 15b=2, then 15c should be recoded to missing. If 15b=3, then 15c should be recoded to missing.

[(c)] "A" before the item number indicates that this item is an optional item from the Appendix. If optional items are used, the NEI-VFQ developers encourage users to use all items for a given sub-scale. This will greatly enhance the comparability of sub-scale scores across studies.

* Response choice "6" indicates that the person does not perform the activity because of non-vision related problems. If this choice is selected, the item is coded as "missing."

Table 3. Step 2: Averaging of Items to Generate VFQ-25 Sub-Scales

Scale	Number of Items	Items to be averaged (after recoding per table 2)
General health	1	1
General Vision	1	2
Ocular Pain	2	4, 19
Near Activities	3	5, 6, 7
Distance Activities	4	8, 9, 14
Vision Specific: Social Functioning Mental Health Role Difficulties Dependency	 2 4 2 3	 11, 13 3, 21, 22, 25 17, 18 20, 23, 24
Driving	3	15c, 16, 16a
Color Vision	1	12
Peripheral Vision	1	10

Table 4. Step 2: Averaging of Items to Generate VFQ-39 Sub-Scales (VFQ-25 + Optional Items)

Scale	Number of Items	Items to be averaged (after recoding per table 2)
General health	2	1, A1
General Vision	2	2, A2
Ocular Pain	2	4, 19
Near Activities	6	5, 6, 7, A3, A4, A5
Distance Activities	6	8, 9, 14, A6, A7, A8
Vision Specific: Social Functioning Mental Health Role Difficulties Dependency	3 5 4 4	11, 13, A9 3, 21, 22, 25, A12 17, 18, A11a, A11b 20, 23, 24, A13
Driving	3	15c, 16, 16a
Color Vision	1	12
Peripheral Vision	1	10

1. Example of VFQ-25 Scoring Algorithm for Near Activities Sub-Scale

5. How much difficulty do you have reading ordinary print in newspapers? Would you say you have:

No difficulty at all 1
A little difficulty 2
Moderate difficulty 3
Extreme difficulty **(4)**
Stopped doing this because of your eyesight 5
Stopped doing this for other reasons or not interested in doing this 6

6. How much difficulty do you have doing work or hobbies that require you to see well up close, such as cooking, sewing, fixing . . . ? Would you say you have:

No difficulty at all **(1)**

A little difficulty 2

Moderate difficulty 3

Extreme difficulty 4

Stopped doing this because of your eyesight 5

Stopped doing this for other reasons or not interested in doing this 6

7. Because of your eyesight, how much difficulty do you have finding something on a crowded shelf? Would you say you have:

No difficulty at all 1

A little difficulty 2

Moderate difficulty 3

Extreme difficulty **(4)**

Stopped doing this because of your eyesight 5

Stopped doing this for other reasons or not interested in doing this 6

Scoring example - Figure 1

Items 5, 6 and 7 are used to generate the near activities sub-scale score (Table 3). Each of the items has 6 response choices. Response choice 6 indicates that the respondent does not perform the activity because of reasons that are unrelated to vision. If a respondent selects this choice, the answer is treated as missing and an average of the remaining items is calculated. Response choice 5 indicates that an activity is so difficult that the participant no longer performs the activity. This extremely poor near vision response choice is recoded to "0" points before taking an average of all three items. To score all items in the same direction, Table 2 shows that responses 1 through 5 for items 5, 6, and 7 should be recoded to values of 100, 75, 50, 25, and 0 respectively. If the respondent is missing one of the items, the person's score will be equal to the average of the two non-missing items.

Formula:

$$Mean = \frac{\text{Score for each item with a non-missing answer}}{\text{Total number of items with non-missing answers}}$$

Example:

$$\text{With responses converted} = \frac{25 + 100 + 25}{3} = 50$$

Note: 100 = Best, 0 = Worst possible score.

Psychometric properties of VFQ-25 sub-scales

Psychometric data for VFQ-25 reported in the earlier pre-publication version of the scoring manual have been updated and submitted for peer-reviewed publication.[2] The values reported in this document are identical to those reported in the future publication and should be used when citing the performance characteristics of the VFQ-25.

Statistical Power Calculations

Tables 8, 9, and 10 are provided to estimate statistical power when using the VFQ-25 and VFQ-39. These tables estimate the number of subjects needed per group to attain 80% power (alpha = 0.05, two-tailed) depending on the anticipated difference in scores between groups. Table 8 contains power calculations for changes over time between two experimental (i.e. randomized) groups using a repeated-measures design. For example, if one were interested in being able to detect a 5-point difference for the VFQ-25 General Vision sub-scale, one would need 271 subjects per group. Table 9 shows power calculations for two

experimental groups using a single, post-intervention measurement design. Such a design is not as precise as a design that uses a baseline and post-intervention measurement points (i.e. more subjects are needed per group to detect the same difference). Table 10 provides corresponding sample size information for a non-experimental (i.e. non- randomized) repeated-measures design where subjects self-select into the two groups. One sees that the number of subjects needed per group is more than that needed for a randomized experiment (Table 8) and less than the number needed for a randomized, post-intervention-only measurement design (Table 9).

Table 8. Sample sizes needed per group to detect differences in change over time between two experimental groups for the VFQ-25, repeated measures design

Number of Points Difference

Scale Name	SD	2	5	10	20
VFQ-25:					
General Health	26.00	1696	271	68	17
General Vision	21.00	1106	177	44	11
Ocular Pain	17.00	725	116	29	7
Near Activities	29.00	2110	338	84	21
Distance Activities	29.00	2110	338	84	21
Social Functioning	27.00	1829	293	73	18
Mental Health	27.00	1829	293	73	18
Role Difficulties	29.00	2110	338	84	21

Dependency	28.00	1967	315	79	20
Driving	35.00	3073	492	123	31
Color Vision	23.00	1327	212	53	13
Peripheral Vision	27.00	1829	293	73	18
VFQ-25 Composite	20.00	1004	161	40	10
VFQ-39:					
General Health	21.00	1106	177	44	11
General Vision	19.00	906	145	36	9
Ocular Pain	17.00	725	116	29	7
Near Activities	28.00	1967	315	79	20
Distance Activities	26.00	1696	271	68	17
Social Functioning	25.00	1568	251	63	16
Mental Health	26.00	1696	271	68	17
Role Difficulties	28.00	1967	315	79	20
Dependency	27.00	1829	293	73	18
Driving	35.00	3073	492	123	31
Color Vision	23.00	1327	212	53	13
Peripheral Vision	27.00	1829	293	73	18
VFQ-39 Composite	21.00	1106	177	44	11

Note: Scales are all scored on 0-100 possible range. Estimates assume alpha = 0.05, two-tailed t-test, power = 80%, and an inter-temporal correlation between scores of 0.60.

Table 9. Sample sizes needed per group to detect differences between two experimental groups for the VFQ-25, post-intervention measures only.

Number of Points Difference

Scale Name	SD	2	5	10	20
VFQ-25:					
General Health	26.00	2650	424	106	26
General Vision	21.00	1729	277	69	17

Ocular Pain	17.00	1133	181	45	11
Near Activities	29.00	3297	527	132	33
Distance Activities	29.00	3297	527	132	33
Social Functioning	27.00	2858	457	114	29
Mental Health	27.00	2858	457	114	29
Role Difficulties	29.00	3297	527	132	33
Dependency	28.00	3073	492	123	31
Driving	35.00	4802	768	192	48
Color Vision	23.00	2074	332	83	21
Peripheral Vision	27.00	2858	457	114	29
VFQ-25 Composite	20.00	1568	251	63	16
VFQ-39:					
General Health	21.00	1729	277	69	17
General Vision	19.00	1415	226	57	14
Ocular Pain	17.00	1133	181	45	11
Near Activities	28.00	3073	492	123	31
Distance Activities	26.00	2650	424	106	26
Social Functioning	25.00	2450	392	98	25
Mental Health	26.00	2650	424	106	26
Role Difficulties	28.00	3073	492	123	31
Dependency	27.00	2858	457	114	29
Driving	35.00	4802	768	192	48
Color Vision	23.00	2074	332	83	21
Peripheral Vision	27.00	2858	457	114	29
VFQ-39 Composite	21.00	1729	277	69	17

Note: Scales are all scored on 0-100 possible range. Estimates assume alpha = 0.05, two-tailed t-test, and power = 80%.

Table 10. Sample sizes needed per group to detect differences between two *self-selected groups* for the VFQ-25, repeated measures design

Number of Points Difference

Scale Name	SD	2	5	10	20
VFQ-25:					
General Health	26.00	2120	339	85	21
General Vision	21.00	1383	221	55	14
Ocular Pain	17.00	906	145	36	9
Near Activities	29.00	2637	422	105	26
Distance Activities	29.00	2637	422	105	26
Social Functioning	27.00	2286	366	91	23
Mental Health	27.00	2286	366	91	23
Role Difficulties	29.00	2637	422	105	26
Dependency	28.00	2459	393	98	25
Driving	35.00	3842	615	154	38
Color Vision	23.00	1659	265	66	17
Peripheral Vision	27.00	2286	366	91	23
VFQ-25 Composite	20.00	1254	201	50	13
VFQ-39:					
General Health	21.00	1383	221	55	14
General Vision	19.00	1132	181	45	11
Ocular Pain	17.00	906	145	36	9
Near Activities	28.00	2459	393	98	25
Distance Activities	26.00	2120	339	85	21
Social Functioning	25.00	1960	314	78	20
Mental Health	26.00	2120	339	85	21
Role Difficulties	28.00	2459	393	98	25
Dependency	27.00	2286	366	91	23
Driving	35.00	3842	615	154	38
Color Vision	23.00	1659	265	66	17
Peripheral	27.00	2286	366	91	23
VFQ-39 Composite	21.00	1383	221	55	14

Note: Scales are all scored on 0-100 possible range. Estimates assume alpha = 0.05, two-tailed t-test, power = 80%, and an inter-temporal correlation between scores of 0.60.

Frequently Asked Questions (FAQ)

Q. What kind of permissions are required to use the VFQ-25 in a research study?

The VFQ-25 is a public document available without charge for all researchers to use provided they identify the measure as such in all publications and cite the appropriate developmental papers. Users do not need to notify the developers or the NEI that they intend to use the measure. However, there are some specific permissions for using the VFQ-25 that are detailed on the cover page of the questionnaire itself. These include acknowledging in all publications that the VFQ-25 was developed by RAND and funded by the NEI, and that any changes made to the measure for your particular study will be identified as such.

Q. Can I change the format of the VFQ-25 to suit my study?

Any change to the wording or order of the items would constitute a change to the measure and should be specified as such in any published papers. Other than this, it is expected that researchers may need to change the format or appearance of items to suit their purposes.

As of August 2000, to our knowledge no studies have reported on the effect of item order on responses to VFQ-25 or other similar vision- targeted surveys. That is, whether responses change depending where particular items appear in the questionnaire. However, to ensure the comparability of scores across studies, it is our position that the order of items should not be changed.

Q. Has the VFQ-25 been translated into any other languages?

As of August 2000, the developers are aware of translation into approximately 9 languages. For the cost of distribution, a Spanish language version for Mexican-American populations is available from the UCLA and RAND based developers. The developers will provide researchers with the names of other persons to contact for other language translations. Should researchers wish to translate the VFQ-25, the same permissions apply, with the additional requirement that all publications specify responsibility for the translation along with instructions for obtaining a copy of the translated version.

Q. Do you have any additional normative information for specific populations?

The developers currently are not conducting studies for the express purpose of further investigating the psychometric properties of the VFQ-25 or producing normative data. However, many researchers are currently using the VFQ-25 as an endpoint or outcome in a number of health services and clinical studies. It is likely that as these studies are completed, results that are relevant to better understanding the performance of the VFQ-25 will accompany the main results of each study. The developers and staff at the NEI are aware of other researchers who are collecting condition-specific normative data on population-based samples with the VFQ-25 and when possible will provide contact information for these investigators to new users.

Q. How relevant is the normative data provided in the scoring manual to my sample?

The means, standard deviations, and statistical power values shown in this document were estimated using cross-sectional data from the Field Test Study. Participants recruited for the Field Test were not randomly sampled, but

rather were identified for enrollment based on clinical criteria biased toward persons with moderate to severe forms of each target disease. Further, because it was our desire to enroll a broad spectrum of patients based on disease severity, we did not take into consideration treatment status. Please see references #3 for a full description of the NEI-VFQ field test study sample.

Q. Why is a single-item general health item included in the VFQ-25?

During the developmental phase of the NEI- VFQ, vision-targeted health-related quality of life (HRQOL) was a relatively new concept. For this reason, we included this question to ensure that researchers had a minimal amount of information about a person's general health status to use as a benchmark against other published samples or cohorts.

This general health rating question has been widely used in studies and is a robust predictor of future health and mortality. However, to fully measure generic HRQOL, many quality of life measurement experts recommend including a separate generic measure of HRQOL such as the SF-36 or SF-12.[9] In such a situation the single- item VFQ-25 general health rating question is not needed because the identical question is asked as part of these surveys.[10,11]

Q. Should we be looking at the sub-scales or the composite score?

The VFQ-25 sub-scales are grouped by theme or domain. So, for example, items having to do with near vision are differentiated from items having to do with other vision activities like distance vision or ocular pain. This does not mean that the items are not highly correlated or that they are psychometrically distinct. What it does mean is that researchers should beforehand carefully consider which vision-specific domains are most likely to be influenced by a particular disease and/or treatment and then focus on the results from those sub-scales to support their findings.

The composite score is best used in situations where an overall measure of vision-targeted health related quality of life is desired. For example, in studies where it is not clear what the specific impact of ocular disease or a new treatment might be. Also, in situations where differences can be hypothesized between groups beforehand across multiple sub-scales but the overall sample size of the study is relatively small, because it is likely that the error term for the composite score is likely to be smaller than for any given sub-scale, it may be more efficient to represent these differences as a single score.

Q. What benefit is there to using the VFQ-25 over a measure more specific to a particular disease, like the Activity of Daily Vision Scale (ADVS)[10] for persons with age-related cataracts?

The VFQ-25 contains items that are very similar to items found in other vision-targeted measure like the ADVS that are more task oriented. However, whereas the ADVS was designed specifically to assess a set of activities most relevant to patients undergoing cataract surgery, the VFQ-25 expands the range of activities to measure the impact of ocular disease on broader domains of health such as social and emotional well-being. Serious ocular diseases that lead to irreversible loss of vision are likely to impact dimensions of a person's life beyond simple tasks such as driving or reading the newspaper, and similarly, by preserving vision, many successful interventions also will impact persons' lives at this more global level. Especially in these situations, use of the VFQ-25 should be considered.

Q. Why does the response to item 15b, "stopped driving due to vision and other reasons", generate a missing score for the subsequent driving items?

Driving items 15, 15a, and 15b are filter questions designed to specify whether a person has ever driven a car, and if so, whether they are currently driving or if they have stopped. If people have never driven a car, then, of course, their answers should be set to missing for all driving items. Similarly, this also applies to people who have stopped

driving for other reasons not due to vision. However, in the course of pilot testing the field test participants wanted this additional mixed response option. It was our decision that although persons did indeed report not driving due to vision, it was not clear how much of a role the "other" reason also played in this decision. Therefore, we set the scoring criteria for this response to be missing for all subsequent driving items to be absolutely sure that all driving responses reflected only problems with vision. Should researchers wish to change this response option to allow persons to answer subsequent driving items (currently there is a skip to item #17), this change should be noted in subsequent publications.

References

1. Mangione CM, Berry S, Lee PP, et al. Identifying the content area for the National Eye Institute Vision Function Questionnaire (NEI-VFQ): Results from focus groups with visually impaired persons. *Arch Ophthalmol.* 1998;116:227-238.
2. Mangione CM, Lee PP, Gutierrez PR, et al. Development of the 25-item National Eye Institute Visual Function Questionnaire (VFQ-25). *Arch Ophthalmol.* 2001 (in press).
3. Mangione CM, Lee PP, Pitts J, et al. Psychometric properties of the National Eye Institute Visual Function Questionnaire, the NEI-VFQ. *Arch Ophthalmol*. 1998;116:1496-1504
4. Gutierrez P, Wilson MR, Johnson C, Gordon M, Cioffi GA, Ritch R; Sherwood M, Meng K, Mangione CM. Influence of glaucomatous visual field loss on health- related quality of life. *Arch Ophthalmol*. 1997;115:777-84.
5. Parrish RK 2nd, Gedde SJ, Scott IU, Feuer WJ, Schiffman JC, Mangione CM, Montenegro-Piniella A. Visual function and quality of life among patients with glaucoma. *Arch Ophthalmol*. 1997;115:1447-55.
6. Quality of life assessment in the collaborative ocular melanoma study: design and methods. COMS-QOLS Report No. 1. COMS Quality of Life Study Group. *Ophthal. Epidemiology*. 1999;6:5-17.
7. Scott IU, Smiddy WE, Schiffman J, Feuer WJ, Pappas CJ. Quality of life of low- vision patients and the impact of low- vision services. *Amer. J. Ophthalmol*. 1999;128:54-62.
8. Cole SR, Beck RW, Moke PS, Gal RL, Long DT. The National Eye Institute Visual Function Questionnaire: experience of the ONTT. Optic Neuritis Treatment Trial. *Invest Ophthalmol Vis Sci* 2000;41:1017-21.
9. Mangione CM, Lee PP, Hays RD. Measurement of visual functioning and health-related quality of life in eye disease and cataract surgery. In B. Spilker (ed.), Quality of Life and Pharmacoeconomics in clinical trials, 2nd edition. New York: Raven Press 1996:1045-1051
10. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-item Health Survey 1.0. *Health Econ* 1993;2:217-227.
11. Ware J Jr, Kosinski M, Keller SD 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220-33A.

APPENDIX 4. Scoring rules for multi-item scales in SF-36 (Ware 2007)

Response scores to different items in the SF-36 have different ranges of pre-coded numeric values. For some items, higher pre-coded values represent a more favorable health state (e.g. '1' corresponds to "Yes, Limited a Lot" for item 3a and '3' corresponds to "No, Not Limited at All"). However, for other items, higher pre-coded values represent worse health (e.g. '1' for item 1 corresponds to "Excellent" and '5' corresponds to "Poor"). To account for these differences in valence, raw pre-coded numeric item scores will be converted to final scores for which higher values always represent better health. Conversion rules are set forth in Table 3 below.

Table 3 Conversion of SF-36 (v2) Raw Item Score to Final Item Scores

Scale	Pre-coded Item Value					
Item Number	1	2	3	4	5	6
Physical Functioning						
3a through 3j	1	2	3	N/A	N/A	N/A
Role Physical						
4a through 4d	1	2	3	4	5	N/A
Bodily Pain						
7	6	5.4	4.2	3.1	2.2	1
8, if item 7 is non-missing and pre-coded value = 1	6	4	3	2	1	N/A
8, if item 7 is non-missing and pre-coded value = 2 through 6	5	4	3	2	1	N/A
8, if item 7 is missing	6	4.75	3.5	2.25	1	N/A
General Health						
1	5	4.4	3.4	2	1	N/A
11a and 11c	1	2	3	4	5	N/A
11b and 11d	5	4	3	2	1	N/A
Vitality						
9a and 9e	5	4	3	2	1	N/A
9g and 9i	1	2	3	4	5	N/A
Social Functioning						
6	5	4	3	2	1	N/A
10	1	2	3	4	5	N/A

Role Emotional 5a through 5c	1	2	3	4	5	N/A
Scale	Pre-coded Item Value					
Item Number	1	2	3	4	5	6
Mental Health						
9b, 9c, and 9f	1	2	3	4	5	N/A
9d and 9h	5	4	3	2	1	N/A
Health Transition						
2	1	2	3	4	5	N/A

If one or more items in a given scale are missing for a given subject, responses will be imputed using the average value for that subject for all non-missing items in the scale, provided that 50% or fewer of the items in the scale are missing. For example, if a subject does not answer one item in the 5-item MH scale, the subject's average score across the four non-missing items will be used to impute the value for the one missing item. If more than 50% of the items in a multi-item scale are missing, however, the entire scale will be set to missing. For single-item measures, the final score will be set to missing if the item raw score is missing.

After conversion of raw item score to final item score and imputation of missing data, a raw score will be computed for each scale. This raw score will be the sum of the final scores for all items in that scale. Raw scale scores will then be transformed into standardized scores, with a range from 0 to 100 (Note: raw and transformed scale scores are not calculated for the single-item Health Transition measure)

The lowest possible raw score and the possible raw score range of each multi-item scale are set forth in Table 4.

Table 4 Lowest/ Highest Possible Raw Scale Scores and Scale Raw Score Range

Scale	Lowest and Highest Possible Raw Scale Score	Scale Raw Score Range
Physical Functioning	10, 30	20
Role Physical	4, 20	16
Bodily Pain	2, 12	10
General Health	5, 25	20

Vitality	4, 20	16
Social Functioning	2, 10	8
Role Emotional	3, 15	12
Mental Health	5, 25	20

After transformation of raw scale scores to standardized scale score, the standardized scale score will be normalized as follows:

- PF Z-score (PF_Z) = (Transformed Score - 83.29094) / 23.75883
- RP Z-score (RP_Z) = (Transformed Score - 82.50964) / 25.52028
- BP Z-score (BP_Z) = (Transformed Score - 71.32527) / 23.66224
- GH Z-score (GH_Z) = (Transformed Score - 70.84570) / 20.97821
- VT Z-score (VT_Z) = (Transformed Score - 58.31411) / 20.01923
- SF Z-score (SF_Z) = (Transformed Score - 84.30250) / 22.91921
- RE Z-score (RE_Z) = (Transformed Score - 87.39733) / 21.43778
- MH Z-score (MH_Z) = (Transformed Score - 74.98685) / 17.75604

Scoring Rules for Overall Component

The two overall components will be computed as follows using the scale normalized scores:

- PCS_Z = (PF_Z * 0.42402) + (RP_Z * 0.35119) + (BP_Z * 0.31754) + (GH_Z * 0.24954) +(VT_Z * 0.02877)

+ (SF_Z * (-0.00753)) + (RE_Z * (-0.19206)) + (MH_Z * (-0.22069))

- MCS_Z = (PF_Z * (-0.22999)) + (RP_Z * (-0.12329)) + (BP_Z * (-0.09731)) + (GH_Z * (-0.01571)) + (VT_Z

* 0.23534) + (SF_Z * 0.26876) + (RE_Z * 0.43407) + (MH_Z * 0.48581)

Transform composite scores PCS and MCS and 8 Z-score to :

Normalized scores = 50 + (Health Domain Z-score × 10)

- PF Normalized Score = 50+(PF_Z*10)
- RP Normalized Score = 50+(RP_Z*10)
- BP Normalized Score =50+(BP_Z*10)
- GH Normalized Score =50+(GH_Z*10)

- VT Normalized Score = 50+(VT_Z*10)
- SF Normalized Score = 50+(SF_Z*10)
- RE Normalized Score = 50+(RE_Z*10)
- MH Normalized Score = 50+(MH_Z*10)
- PCS Normalized Score = 50+(MCS_Z *10)
- MCS Normalized Score = 50+(PCS _Z*10)