Clinical Development

RTH258/Brolucizumab

CRTH258B2301 / NCT03481634

A Two-Year, Three-Arm, Randomized, Double-Masked, Multicenter, Phase III Study Assessing the Efficacy and Safety of Brolucizumab versus Aflibercept in Adult Patients with Visual Impairment due to Diabetic Macular Edema (KESTREL)

Statistical Analysis Plan (SAP)

Author: [redacted]

Document type: SAP Documentation

Document status: Amendment V1.0 – clean version

Release date: 26-Nov-2020

Number of pages: 51

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
26-Jun-2020	Prior to DB lock	Creation of first version	N/A	NA
26-Nov-2020	Prior to DB lock	Creation of amendment 1	Reflect the change in the analysis timing to occur when all subjects complete Week 52 visits instead of the first 534 randomized subjects	Sections 2.1, 2.13
			Add analyses to cover potential impact of the COVID-19 pandemic	Sections 2.3.1, 2.5.4.2, 2.5.4.3, 2.6.4, 2.7.2.2, 2.7.2.3, 2.8.1
			Add back the summary tables of medical history and prior/concomittant medications, as per the protocol	Sections 2.3.3, 2.4.2
			Add clarifications for the definition of primary and supplementary estimands	Sections 2.5.4.1, 2.5.4.2
			Add clarifications for the q12w treatment status	Section 2.6.2
			Update testing strategy to add confirmatory superiority testing on efficacy endpoints	Section 2.7.2.1
			Add details on the determination of AESIs	Section 2.8.1.1
			Add section regarding additional imaging parameters for subjects with AESI in the study eye	Section 2.8.4.4
			Correct minor issues in the first version	

Table of contents

List of abbreviations

ADA	Anti-drug antibody
AE	Adverse Event
ALP	Alkaline Phosphate
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AR	Analysis Restrictions
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Classification
BCVA	Best Corrected Visual Acuity
CI	Confidence Interval
cm	Centimeter
CF	Color Fundus
COVID-19	Coronavirus Disease 2019
CSFT	Central subfield thickness
CSFTns	Central Subfield Thickness-neurosensory retina
CSR	Clinical Study Report
CRS	Case Retrieval Strategy
CRC	Central Reading Center
DBL	Database Lock
DMC	Data Monitoring Committee
DME	Diabetic Macular Edema
DR	Diabetic Retinopathy
DRSS	Diabetic Retinopathy Severity Scale
eCRF	Electronic Case Report Form
eCRS	Electronic Case Retrieval Strategy
ETDRS	Early Treatment Diabetic Retinopathy Study
EoT	End of Treatment
EoS	End of Study
FA	Fluorescein Angiography
IA	Interim Analysis
IOP	Intraocular Pressure
IRT	Interactive Response Technology
IRF	Intraretinal Fluid
IVT	Intravitreal Treatment
kg	Kilogram
KM	Kaplan Meier
LOCF	Last observation carried forward
LSM	Least Square Means
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligrams
MICE	Multiple Imputation by Chained Wquations

mL	Millilitres
MMRM	Mixed Model for Repeated Measures
OCT	Optical Coherence Tomography
PD	Protocol Deviation
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD-OCT	Spectral Domain Optical Coherence Tomography
SE	Standard Error
SOC	System Organ Class
ULN	Upper Limit of Normal
q1	25th Percentile
q3	75th Percentile
q4w	Every 4 Weeks
q6w	Every 6 Weeks
q8w	Every 8 Weeks
q12w	Every 12 Weeks
SRF	Subretinal Fluid
TBL	Total Bilirubin
TEAE	Treatment-Emergent Adverse Event
TFLs	Tables, Figures, Listings
VA	Visual acuity
VFQ-25	Visual Functioning Questionnaire-25
WHO	World Health Organization

1 Introduction

The purpose of this Statistical Analysis Plan (SAP) is to describe the implementation of statistical analyses planned in the study protocol, and to provide detailed statistical methods that will be used for the Clinical Study Report (CSR) of study CRTH258B2301.

Data will be analyzed according to the data analysis Section 9 of the study protocol which will be available in Appendix 16.1.1 of the CSR. Important information is given in the following sections and additional details will be provided, as applicable, in Appendix 16.1.9 of the CSR.

The SAP will be finalized before the interim database lock (DBL) for the primary analysis at Week 52. Any changes to the SAP after approval will be documented.

The following document was referenced while writing this SAP:

CRTH258B2301 Clinical Trial Protocol Final version 03 dated 12-Jun-2020

1.1 Study design

This is a randomized, double-masked, multi-center, active-controlled 3-arm study in patients with diabetic macular edema (DME) to evaluate the safety and efficacy of brolucizumab 3 mg and 6 mg against the active control aflibercept 2 mg.

Approximately 700 patients will be screened in order to randomize a total of approximately 534 patients (178 per arm) in a 1:1:1 ratio in one of the three treatment arms:

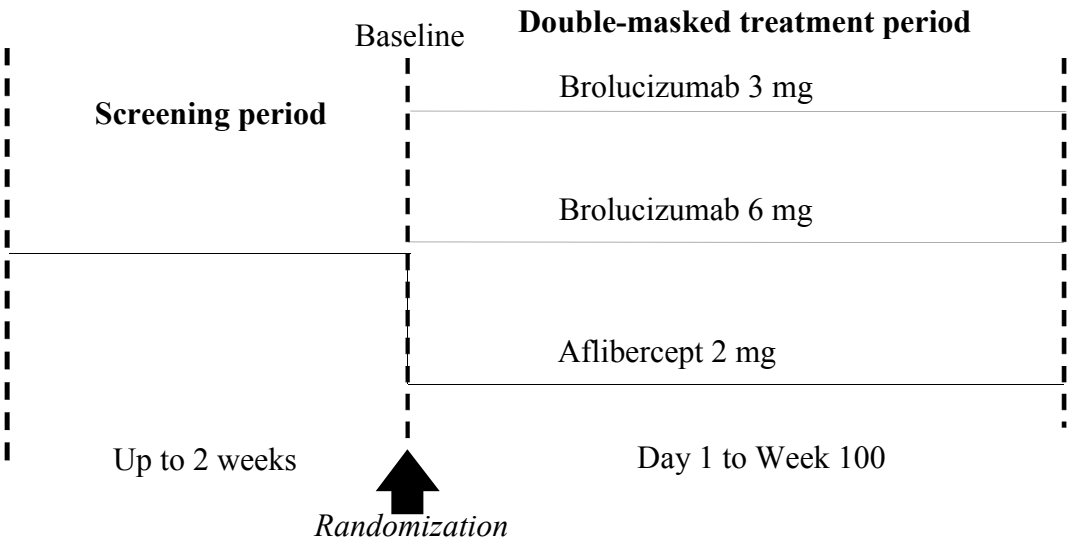
- Brolucizumab 3 mg: 5 × q6w loading then q12w/q8w maintenance
- Brolucizumab 6 mg: 5 × q6w loading then q12w/q8w maintenance
- Aflibercept 2 mg: 5 × q4w loading then q8w maintenance

At the baseline visit, all eligible patients will be randomized via Interactive Response Technology (IRT) to one of the treatment arms. Stratification for Japanese ethnicity (Japanese vs. non-Japanese) will be considered.

Since the treatment schedule is different for brolucizumab and aflibercept treatment arms the following will be applied to ensure masking:

- In addition to visits every 4 weeks for all patients for 2 years, extra visits are scheduled at Weeks 6 and 18 for all treatment arms
- The patients will receive active/sham injection at each protocol visit except at Weeks 20, 28 and 100 visits (no scheduled treatment for any arm)
- Disease activity assessment will be performed for all arms at pre-specified visits
- To fulfil the double-masking requirement, the investigational site will have masked and unmasked staff

Figure 1-1 Study design



There will be two treatment phases for IVT injections with different timing for brolucizumab and aflibercept treatment arms:

Loading Phase:

Brolucizumab 3 mg or 6 mg: In the loading phase, treatment with brolucizumab 3 mg or 6 mg will occur every 6 weeks for five (5) consecutive injections (Baseline, Weeks 6, 12, 18 and 24). To preserve the masking, the patients assigned to this regimen will receive sham injection on Weeks 4, 8 and 16.

Aflibercept 2 mg: In the loading phase, treatment with aflibercept 2 mg will occur every 4 weeks for five (5) consecutive injections (Baseline, Weeks 4, 8, 12 and 16). To preserve the masking, the patients assigned to aflibercept arm will receive sham injection on Weeks 6 and 18.

Maintenance Phase:

The treatment interval during the maintenance phase will be as follows:

Brolucizumab 3 mg or 6 mg: From Week 24 onwards, patients will be scheduled to receive one injection of brolucizumab 3 mg or 6 mg every 12 weeks. If, however, disease activity is identified by the evaluating/masked investigator at pre-specified visits, the patient will be assigned to receive treatment every 8 weeks (please refer to protocol for 'Evaluation of Disease Activity'). A disease activity assessment will also be performed at Week 96 to document the adequacy of the q12w treatment schedule at the end of the 2 year follow-up but will not be entered into IRT and will have no effect on the patient's treatment schedule.

Aflibercept 2 mg: From Week 16 onwards, patients will receive one injection of aflibercept 2 mg every 8 weeks (first injection after Week 16 to be given at Week 24) until Week 96 visit. Disease activity assessments will be conducted at pre-specified visits by the evaluating/masked investigator for masking purposes and will not influence the treatment interval.

The primary analysis will be based on the Week 52 data, i.e. all data up to and including Week 52 (see Section 2.1).

Patients will remain in the study and will continue to receive masked treatment through the planned duration (100 weeks) to allow for further masked evaluation of efficacy and safety. Treatment masking of individual patients will remain intact for all patients, investigators and selected staff from the Sponsor who have contact with patients or investigators or those who are involved in the direct conduct of the study until the final database lock has occurred.

1.2 Study objectives and endpoints

Study objectives and related endpoints are described in Table 1-1 below.

Table 1-1 Objectives and related endpoints

Objective	Endpoint
Primary	
To demonstrate that brolucizumab is noninferior to aflibercept with respect to the visual outcome after the first year of treatment	Change from baseline in best-corrected visual acuity (BCVA) at Week 52
Secondary	
To demonstrate that brolucizumab is noninferior to aflibercept with respect to visual outcome during the last 3 months of the first year of treatment	Change from baseline in BCVA averaged over a period Week 40 to Week 52
To estimate the proportion of patients treated at q12w frequency with brolucizumab	Proportion of patients maintained at q12w up to Weeks 52 & 100
To estimate the predictive value of the first q12w cycle for maintenance of q12w treatment with brolucizumab	Proportion of patients maintained at q12w up to Weeks 52 & 100, within those patients that qualified for q12w at Week 36
To evaluate the functional and anatomical outcome with brolucizumab relative to aflibercept	Change from baseline by visit up to Week 100 in BCVA and in parameters derived from SD-OCT, Color fundus photography and Fluorescein angiography
To evaluate the effect of brolucizumab relative to aflibercept on the Diabetic Retinopathy status	Change in ETDRS Diabetic Retinopathy Severity Scale (DRSS) score up to Week 100
To assess the safety and tolerability of brolucizumab relative to aflibercept	Incidence of Ocular and Non-ocular AEs, vital signs and laboratory values up to Week 100
To evaluate the effect of brolucizumab relative to aflibercept on patient-reported outcomes (VFQ-25)	Change in patient reported outcomes (VFQ-25) total and subscale scores from baseline up to Week 100

2 Statistical methods

2.1 Data analysis general information

The primary safety and efficacy analysis will be based on the Week 52 data, i.e. all data up to and including Week 52. This analysis will be performed when all randomized subjects have completed their Week 52 visit or terminated the study before Week 52, while subjects continue to receive masked treatment through the planned study duration of 100 weeks.

A second planned interim analysis may be performed by locking the Week 76 data in case of regulatory request of supplemental data to be submitted during the review period.

The statistical analysis will be performed by Novartis using SAS Version 9.4 or above.

Continuous variables will be summarized using the number of observations, mean, standard deviation, standard errors (SE), median, quartiles, minimum and maximum values. Categorical variables will be summarized with number of observations, the number of observations for each category and the corresponding percentage. Where appropriate, 2-sided 95% confidence intervals (CIs) for point estimates of the mean or proportion will be provided. For the treatment difference brolucizumab – aflibercept, point estimates, 95% CIs will be provided as appropriate unless otherwise specified.

All the analyses listed in the SAP that correspond to data collected during the 2nd year of the study (post-Week 52) will be part of the end of study (year-2) CSR only. For the year-2 CSR, efficacy endpoints specific to the first year, i.e., up to the Week 52 visit (e.g., the change from baseline in BCVA averaged over Week 40 to Week 52) will not be analyzed and reported again; analyses by visit and assessments based on cumulative data (e.g., incidence of AEs) will include data from baseline up to the end of study.

2.1.1 General definitions

Study drug and study treatment

Study drug refers to brolucizumab 3 mg, brolucizumab 6 mg and aflibercept 2 mg.

Study treatment refers to study drug or sham IVT injections.

Study day

Day 1 is defined as the date of first dose of study drug (brolucizumab or aflibercept). Study day is defined as the number of days relative to the date of first dose of study treatment (Day 1).

Therefore, for a particular date, study day will be calculated as follows:

- for dates on or after the date of first administration of study treatment:

  Study day = Assessment date – Date of first dose of study treatment + 1;
- for dates prior to the date of first administration of study treatment:

  Study day = Assessment date – Date of first dose of study treatment.

Baseline

The baseline value is defined as the last assessment performed prior to administration of the first dose of study treatment.

All data collected after first study treatment are defined as post-baseline.

End of study day mapping

The end of study (EoS) date is the date when a subject completes or discontinues the study.

For reporting data by visit in outputs, the end of study visit will be allocated to the actual (reported) visit number. If end of study date is not on a scheduled visit, then the EoS visit will be allocated, based on study day, to the closest future scheduled study visit.

End of treatment day mapping

The "Date of Last Exposure" is the date of the last study treatment on or prior to the end of treatment (EoT) date.

For reporting data by visit in outputs, the EoT visit will be allocated to the actual (reported) visit number. If EoT date is not on a scheduled visit, then the EoT visit will be allocated, based on study day, to the closest future scheduled study visit.

Unscheduled visits

Data collected at unscheduled visits will not be used in 'by-visit' tabulations or graphs , but will be included in analyses based on all post-baseline values such as last observation carried forward (LOCF) imputation, average BCVA change from baseline over a given period, and summary of maximum letter loss in BCVA from baseline. These data would not be used in case of analyses with mixed model for repeated measures (MMRM).

All data collected at unscheduled visits will be included in listings.

Missing and implausible dates

The general approach to handling missing dates is shown in Section 5.1.

2.2 Analysis sets

The **All Enrolled Set** includes all subjects who signed informed consent. This analysis set will be used to summarize subject disposition.

The **Randomized Set** will consist of all randomized subjects. Subjects are considered randomized when they have been deemed eligible for randomization by the investigator and given a randomization number. Subjects will be analyzed according to the treatment assigned to at randomization.

The **Full Analysis Set** (FAS) includes all randomized subjects who receive at least one IVT injection of the study treatment. The full analysis set will serve as the primary analysis set for all efficacy analyses. Subjects will be analyzed according to the treatment assigned to at randomization.

Supportive analyses of the primary and key secondary endpoints will include analyses using the **Per Protocol Set** (PPS). PPS is a subset of the FAS and will exclude or censor subjects with important protocol deviations (PDs) and analysis restrictions (ARs) that are expected to majorly affect the validity of the assessment of efficacy and/or safety at Week 52, including for e.g. lack

of compliance (including missed treatments and treatment misallocation), missing data, prohibited concomitant medication and deviations from inclusion/exclusion criteria. Confounded data or discontinuation from study treatment due to lack of efficacy and/or safety do not constitute a reason for exclusion from the PPS.

Before the Week 52 database lock the relevant protocol deviations will be identified at the patient level in the database. After the Week 52 database lock, analysis restrictions will be derived in the analysis database. Censoring applied in relation to the specific PDs / ARs will be specified as well.

The FAS will be the analysis set for the primary estimand as defined in Table 2-1. However, when assessing the robustness of the overall efficacy conclusions, considerations will be given to the analysis based on the primary estimand using FAS and the supplementary estimand (see Table 2-1) using PPS, i.e., similar conclusions on non-inferiority based on both estimands are expected. Inconsistencies in key efficacy study results between the FAS and PPS will be examined and discussed in the clinical study report (CSR).

The **Safety Analysis Set** (SAF) will include all subjects who receive at least one IVT injection. Subjects in the safety analysis set will be analyzed according to the treatment arm from which they received the majority of treatments up to and including Week 48.

Prior to locking the database for the primary analysis at Week 52 and breaking the masked treatment assignment code, the relevant important protocol deviations will be identified as specified in Section 5.6. The corresponding identifications at the subject level including data exclusion from PPS and censoring will be captured in the database. Analysis Restrictions (non protocol deviations) will be identified by programming (as specified in the programming specification document) independently to the treatment arm.

Rules of exclusion criteria of analysis sets are in Appendix Section 5.5.

For the primary analysis performed when the first 534 randomized subjects have completed their Week 52 visit or terminated the study before Week 52, footnotes will clarify that the analysis sets considered for the outputs are not considering all randomized subjects.

2.2.1 Subgroups of interest

The subgroups of interest are specified below:

- Age category (<65, ≥65 years)
- Gender (male, female)
- Diabetes type (Type 1, Type 2)
- Baseline HbA1c (<7.5, ≥7.5%)
- Baseline BCVA categories (≤65, >65 letters)
- Duration of DME since the primary diagnosis (≤3, >3-<12, ≥12 months)
- DME type (focal, diffuse) as per central reading center (CRC)
- Baseline central subfield thickness (CSFT) (<450, ≥450-<650, ≥650 µm)
- Baseline status of intraretinal fluid (IRF) (presence, absence)
- Baseline status of subretinal fluid (SRF) (presence, absence)
- Ethnicity (Japan, non-Japan)

Subgroup analyses will be performed for the primary and key secondary efficacy variables only (as defined in Section 2.5.1 and Section 2.6.1), using the primary analysis approach. More details can be found in Section 2.5.4 and Section 2.6.2.

2.3 Subject disposition, demographics and other baseline characteristics

2.3.1 Subject disposition

The following summaries will be included in the disposition table considering all enrolled subjects: number and percent of subjects who were enrolled into the study, treated, completed the study (Week 52/Week 100), discontinued the study (prior to or at Week 52/Week 100) (including reasons for discontinuation) and discontinued from study treatment (prior to or at Week 52/Week 96) (including reasons for discontinuation).

The number and percentage of subjects who discontinued the study and who discontinued study treatment will be presented by treatment arm and study visit. The number and percentage of subjects treated by site and treatment arm will be presented.

A listing of subjects who discontinued from the study and/or treatment early will be provided by treatment arm. The listing will identify the visits completed and when the study or treatment was discontinued including the corresponding reasons.

Subjects who signed an informed consent form and who were subsequently found to be ineligible prior to randomization will be considered a screen failure. Screen failure information will not be summarized but only listed.

Number and percentage of subjects who were excluded (i.e. not evaluable) from each of the SAF, FAS, and PPS will be presented using the randomized analysis set. A listing of subjects along with the analysis set that they were excluded from and the corresponding reasons will also be presented.

Number and percentage of subjects with important protocol deviations (PD) and analysis restrictions (AR) will be presented by treatment arm and deviation/restriction category. Due to the COVID-19 pandemic, higher number of PDs are expected. To evaluate the PDs that occurred due to COVID-19, the number and percentage of subjects with PDs that occurred due to COVID-19 outbreak will also be provided by deviation category and treatment arm. A listing of all ARs and PDs will be provided by treatment arm and subject, including the information if the AR/PD leads to the subject exclusion from an analysis set and the relationship to COVID-19.

2.3.2 Demographics and baseline characteristics

Demographics and baseline characteristics will be summarized with descriptive statistics for the FAS by treatment arm and overall. Demographic characteristics will include age, gender, race, and ethnicity. The summary of baseline ocular characteristics will be presented for the study eye only and listed separately for the study eye and the fellow eye.
Ocular baseline characteristics include:

- Study eye selection (left eye OS or right eye OD),
- Diabetes type (Type 1, Type 2),

- Duration of DME since the primary diagnosis as a continuous variable and using categories (<3, >3-<12, ≥12 months),
- Macular edema type (focal, diffuse) as per CRC,
- Baseline BCVA as continuous variable and using categories (≤65, >65 letters, and <60, ≥60-≤70, >70 letters),
- Baseline HbA1c as continuous variable and using categories (<7.5, ≥7.5%),
- Baseline CSFT (<450, ≥450-<650, ≥650 µm),
- Baseline status of IRF (presence, absence),
- Baseline status of SRF (presence, absence),
- Baseline ETDRS DRSS score using categories (12-point scale, see Table 2-4).

Duration of DME since diagnosis (months) will be derived as [(first dose date – diagnosis start date + 1)/(365.25/12)]. In case of partial dates, the imputation rule is specified in Section 5.1.4.

Other relevant baseline information will be listed and summarized with descriptive statistics as appropriate.

No tests for differences in demographics and baseline characteristics between treatment arms will be performed. Potential differences will be assessed based on clinical relevance.

2.3.3 Medical history

Medical history and current medical conditions will be summarized and listed for ocular (study eye) and non-ocular events.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment exposure

Extent of exposure to study treatment is calculated as the number of study treatment injections received.

Descriptive statistics for exposure to study treatment will be provided for the Safety set.

The following summaries will be presented:

- Overall number of treatments cumulatively for the period baseline to Week 48/52 (Week 96) including separate analysis for the loading phases, i.e. up to Week 16 (last treatment of the 5 x q4w loading of aflibercept) and up to Week 24 (last treatment of 5 x q6w loading of brolucizumab), and maintenance phase, using the following categories: active and sham IVT injections, active only, sham only
- Treatment exposure by visit: the number and percentage of subjects who received active IVT injections of study treatment, sham injections, missed a treatment (active and sham) and missed visits will be presented by treatment arm and visit
- Frequency of all observed dosing patterns from baseline to Week 52 (Week 100), differentiating between active and sham treatments, missed study treatments and wrong study treatments

- Brolucizumab treatment allocation by visit from Week 32 onwards: number and percentage of subjects on q12w and q8w at each visit, including number of subjects switched from q12w to q8w.

Exposure data will be listed for all treatment arms.

2.4.2 Prior and concomitant medications

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Concomitant medications are defined as medications received after the start of study treatment including those already started prior to the start of the study treatment.

Prior and concomitant medications will be coded according to the WHO Drug Reference List dictionary, with Anatomical Therapeutic Classification (ATC) class and preferred term.

Ocular and non-ocular prior and concomitant medications will be summarized and listed by ATC class and preferred term (PT) by treatment arm. Ocular medications will be listed for the study eye and the fellow eye separately.

Anti-VEGF medications will be summarized by ATC class and preferred term for systemic route, the study eye and the fellow eye separately by treatment arm.

Ocular concomitant non-drug therapies and procedures will be summarized for the study eye only. Both ocular and non-ocular concomitant non-drug therapies and procedures will be listed.

In the summary tables, data collected after the subject discontinued study treatment and started alternative DME treatment in the study eye will be censored (from the day the subject started alternative DME treatment onwards).

2.5 Analysis of the primary and first key secondary endpoints

2.5.1 Primary and first key secondary endpoints

The primary endpoint is the change from baseline in BCVA at Week 52 in the study eye (ETDRS letters).

The first key secondary endpoint is the average change from baseline in BCVA over the period Week 40 through Week 52 in the study eye. For each subject, this endpoint is defined as the average of the changes from baseline to Weeks 40, 44, 48 and 52.

The motivation for the choice of this endpoint is that, averaging the BCVA values over Week 40 to Week 52 will address both random fluctuations and potential trough and peak values during the different treatment cycles (q8w and q12w).

The primary analysis of the primary and first key secondary endpoints will be based on the FAS with last observation carried forward (LOCF) imputation of missing or censored BCVA values.

The primary estimand for the primary endpoint includes the following components:

- Population: Subjects with visual impairment due to DME as per the inclusion/exclusion criteria

- Endpoint: The primary endpoint is the change from baseline in BCVA at Week 52. BCVA will be assessed by the masked investigator using ETDRS-like charts at an initial distance of 4 meters.
- Treatment of interest: The randomized study treatment (brolucizumab or aflibercept)
- The handling of the remaining intercurrent events as follows:
  - Study discontinuation due to any reason: data imputed with LOCF
  - Treatment discontinuation due to any reason: use all the data
  - Data after the start of alternative DME treatment will be censored
  - Summary measure: Difference in the change from baseline in BCVA at Week 52 between brolucizumab and aflibercept treatment arms.

The primary estimand for the first key secondary endpoint has similar components.

2.5.2 Statistical hypothesis, model, and method of analysis

The objective related to the primary and first key secondary endpoints is to demonstrate non-inferiority of brolucizumab to aflibercept with respect to change from baseline in BCVA, considering a margin of 4 ETDRS letters.

Let:

B3 = Brolucizumab 3 mg - 5 x q6w loading then q12w/q8w maintenance

B6 = Brolucizumab 6 mg - 5 x q6w loading then q12w/q8w maintenance

A = Aflibercept 2 mg - 5 x q4w loading then q8w maintenance

The following non-inferiority hypotheses are related to a non-inferiority margin of 4 letters:

$\mathbf{H_{01}}$: $\mu_{B6} - \mu_A \leq$ -4 letters vs. $\mathbf{H_{A1}}$: $\mu_{B6} - \mu_A >$ -4 letters

$\mathbf{H_{02}}$: $\varphi_{B6} - \varphi_A \leq$ -4 letters vs. $\mathbf{H_{A2}}$: $\varphi_{B6} - \varphi_A >$ -4 letters

$\mathbf{H_{03}}$: $\mu_{B3} - \mu_A \leq$ -4 letters vs. $\mathbf{H_{A3}}$: $\mu_{B3} - \mu_A >$ -4 letters

$\mathbf{H_{04}}$: $\varphi_{B3} - \varphi_A \leq$ -4 letters vs. $\mathbf{H_{A4}}$: $\varphi_{B3} - \varphi_A >$ -4 letters

where μ_{B6} μ_{B3} and μ_A are the corresponding unknown true mean changes from baseline in BCVA at Week 52; φ_{B6}, φ_{B3} and φ_A are the corresponding unknown true mean changes from baseline in BCVA averaged over the period Week 40 to Week 52.

Based on the FAS, the above hypotheses will be tested via an analysis of variance (ANOVA) model. The model will include treatment, baseline BCVA (≤65, >65 letters) and age category (<65, ≥65 years) as factors. Two-sided 95% confidence interval (CI) for the least square mean (LSM) difference (brolucizumab - aflibercept) will be presented in letters. Non-inferiority will be considered established if the lower limit of the corresponding 95% CI is greater than -4 letters. P-value for treatment comparison (2-sided) and p-value for non-inferiority (4 letter margin) (1--sided) will be presented.

These 4 hypotheses will be tested sequentially in the order of their numbering (H_{An}, n=1, 2, 3, 4), i.e., confirmatory testing of the second, third or fourth hypotheses requires rejection of each preceding null hypothesis. In this setting, each hypothesis will be assessed at a one-sided significance level of 0.025.

The primary estimand and other supplementary estimands of interest are described in Table 2-1 below, together with their key attributes. The same approach for non-inferiority assessment in change from baseline in BCVA at Week 52 and average change from baseline in BCVA over the period Week 40 through Week 52 will be applied to any supplementary estimand.

Table 2-1 Primary and supplementary estimands

Estimand	Estimand definition	Analysis set	Statistical methods (Including strategy for imputation/replacement of missing/censored data)
Primary estimand	Difference in change from baseline in BCVA at Week 52 excluding the effect of switching to alternative DME medication in the study eye	FAS	Analysis of variance (ANOVA) model including terms for treatment, baseline BCVA (≤65, >65 letters) and age category (<65, ≥65 years), and using LOCF imputation/replacement for missing/censored data.
Supplementary estimand	Difference in change from baseline in BCVA at Week 52 for subjects adhering to the protocol as per the PPS definition	PPS	ANOVA model as per the primary estimand. LOCF imputation/replacement for missing/censored data

For additional information on data handling related to intercurrent events, see Section 5.5 and Section 5.6.

2.5.3 Handling of missing values/censoring/discontinuations

Missing BCVA values will be imputed by LOCF as a primary approach. Observed values from both scheduled and unscheduled visits will be used for the LOCF imputation. For subjects with no post-baseline BCVA value, the baseline value will be carried forward.

For subjects who discontinue study treatment but continue in the study, data collected after the switch to alternative DME treatment in the study eye (see Table 2-2) will be censored for the primary analysis. Censored data will be replaced using LOCF with the last observation collected prior to the start of alternative DME treatment in the study eye.

Table 2-2 Potential alternative DME treatment in the study eye

- Ranibizumab
- Aflibercept
- Bevacizumab (off-label use)
- Laser photocoagulation, previous standard of care still being used as mono- or combination therapy with anti-VEGF
- Intraocularly administered steroids:
  - Dexamethasone
  - Fluocinolone acetonide

From an estimand perspective, the main focus is to adequately reflect in the analysis unfavorable study outcomes related to the treatment (e.g. lack of efficacy, safety problems).

The LOCF approach is expected to be sensitive to an early study termination due to lack of efficacy assuming that such lack of efficacy is reflected in the last observed BCVA measurement. In case of the use of alternative treatment for the underlying disease (DME), data collected after the start of such a treatment would be censored. LOCF will then be based on the last value prior to the start of this treatment, again expecting that this value would reflect the negative BCVA outcome under study treatment. In case of missing data due to lack of safety/tolerability with impairment of the function of the study eye the LOCF method would also provide a sensitive approach to capture such an unfavorable outcome.

In case of missing data occurring independently of the response to study treatment, the LOCF approach assumes stability which seems to be adequate based on historical data both for the maintenance treatment phase (i.e. stabilization of BCVA) and also in case of the absence of any treatment effect with an average natural disease progression in terms of BCVA of only 1-2-letter loss over 1 year. In case of an early study termination during the loading phase occurring independently of the response to study treatment, the LOCF method will result in a conservative estimate potentially underestimating the true outcome.

LOCF is an established method within the assessment of efficacy of anti-VEGF treatments in terms of BCVA outcome. Non-inferiority studies should follow the main design features (primary variables, the dose of the active comparator, eligibility criteria, etc.) as the previously conducted superiority trials in which the active comparator demonstrated clinically relevant efficacy. The primary endpoint in aflibercept Phase III studies VIVID and VISTA was the BCVA change from Baseline to Week 52 with missing data imputed based on LOCF. Based on those studies, the percentage of missing data regarding BCVA is not considered critical (<10%) which limits the impact of the missing data imputation method.

2.5.4 Sensitivity and supportive analyses

2.5.4.1 Sensitivity analyses on the primary estimand

Sensitivity to the statistical model and imputation assumptions from the primary estimand will be considered, using the primary analysis set (FAS) only.

An alternative method of handling missing/censored values as described below may be considered to assess the robustness of the hypothesis testing resulting from the primary analysis described in Section 2.5.2:

- Mixed model for repeated measures (MMRM) assuming missing at random (MAR) using observed data only (including censoring of BCVA values collected after the start of alternative DME treatment). The MMRM will include treatment, visit, baseline BCVA category, age category and treatment by visit interaction as fixed-effect terms, and visit as a repeated measure. An unstructured covariance matrix will be used to model the within-subject error. For the MMRM analysis:
  - The treatment difference brolucizumab - aflibercept at Week 52 will be estimated using the LSM and 95% CI.

- For the endpoint of average change from baseline over the period Week 40 through Week 52, a SAS code using the ESTIMATE statement in PROC MIXED will be provided in the programming specification document to obtain the LSM estimate and 95% CI for the corresponding treatment difference.
- If an MMRM model with an unstructured covariance matrix does not converge, a more restricted covariance matrix can be considered in the following order until convergence is reached: compound symmetry (CS), first-order autoregressive (AR), Toeplitz (TOEP), and variance components (VC).

In this analysis, data collected after the switch to alternative DME treatment in the study eye will be censored.

Other sensitivity analyses on the primary estimand might be considered, such as tipping point analysis or multiple imputation by chained equations (MICE) method.

2.5.4.2 Supportive analysis using a supplementary estimand

Supplementary estimand on the PPS:

The target population, the primary endpoint, the treatment of interest and the summary measure of the supplementary estimand are the same as for the primary estimand. The handling of the intercurrent events for the supplementary estimand can be found in Table 5-6 for the PPS population.

The supportive analysis on the supplementary estimand will apply the same LOCF/ANOVA method as for the primary estimand.

Supplementary estimand to assess the impact of COVID-19:

Another supplementary estimand might be defined to assess the impact of intercurrent events associated with study treatment discontinuation due to COVID-19 on the study conclusions. For subjects who discontinue study treatment due to COVID-19 but continue in the study, data collected after the treatment discontinuation will be censored for the analysis. Censored data will be replaced using LOCF with the last observation collected prior to the study treatment discontinuation. This analysis will be conducted on the FAS if at least 5% of subjects discontinued treatment due to COVID-19.

2.5.4.3 Summary statistics and subgroup analysis

Summary statistics:

- Descriptive statistics of BCVA primary and first key secondary endpoints will use observed data and primary analysis set (FAS), with and without censoring data after use of alternative DME treatment in the study eye.

Subgroup analyses will be conducted to assess the consistency of treatment effect across various subgroups of interest as described in Section 2.2.1. They will be conducted using the framework for the primary estimand only (FAS with censoring of data collected after use of alternative DME treatment in the study eye, and missing/censored values imputed/replaced using LOCF):

- Subgroup analyses will be conducted using the same model and analysis strategies described for the primary and first key secondary endpoints but fitted by category of each of the subgroups. Subgroup variables that are used as fixed effects in the model will be removed from the model statement for the subgroup analysis.
- In case of analyses on subgroups with extremely imbalanced sample sizes, the subgroup levels can either be combined, if appropriate, or the extremely small subgroup will be excluded while fitting the analysis model.
- The point estimate and 95% CI for the between treatment difference for each subgroup will be presented using forest plots.

Subgroup analyses to evaluate impact of COVID-19 pandemic:
As per internal guidance, a sensitivity analysis related to the exposure of subjects to COVID-19 will be conducted. The definition of start and end dates by geographical areas to be used for the sensitivity analysis are as indicated below:

Region/Country	Start Date	End Date
China	01-Jan-2020	End date has not yet been defined
South Korea	20-Feb-2020	End date has not yet been defined
Japan	21-Feb-2020	End date has not yet been defined
Italy	23-Feb-2020	End date has not yet been defined
Rest of the World	01-Mar-2020	End dates have not yet been defined

Non-exposed subjects to COVID-19 are defined as subjects who:
- completed Week 52 visit prior to the pandemic start date,
- or withdrew the study prior to the pandemic start date,
- or withdrew treatment and started alternative DME treatment prior to the pandemic start date.

Exposed subjects to COVID-19 are therefore defined as subjects who:
- did not complete Week 52 visit prior to the pandemic start date (while remaining in the study at the time of pandemic start date),
- or withdrew the study on or after the pandemic start date,
- or withdrew treatment and started alternative DME treatment on or after the pandemic start date,
- or withdrew treatment before the pandemic start date and did not start alternative DME treatment before the pandemic start date, while remaining in the study.

Furthermore, impacted subjects to COVID-19 pandemic were defined as subjects who:
- were exposed to COVID-19 as per the above definition,
- and missed at least one active injection due to COVID-19.

Non-impacted subjects to COVID-19 are therefore defined as subjects who:
- were not exposed to COVID-19 as per the above definition,
- or were exposed to COVID-19 but did not miss any active injection due to COVID-19.

Subgroup analyses will be conducted using the same model and analysis strategies described for the primary and first key secondary endpoints in the exposed and non-exposed subgroups, and in the impacted and non-impacted subgroups. In addition, demographics and baseline

characteristics will be summarized for exposed and non-exposed subjects, and for impacted and non-impacted subjects.

2.6 Analysis of additional key secondary endpoints

2.6.1 Additional key secondary endpoints

Additional key secondary endpoints are:

- Proportion of subjects maintained at q12w up to Week 52 (for brolucizumab treatment arms only)
- Proportion of subjects maintained at q12w up to Week 52, within those subjects that qualified for q12w at Week 36 (for brolucizumab treatment arms only)

2.6.2 Statistical hypothesis, model, and method of analysis

No hypothesis will be tested for the additional key secondary efficacy endpoints.

Following the estimand concept, in which consequences of lack of efficacy and/or lack of safety need adequate reflection in the efficacy estimates, the primary approach to derive the proportion of subjects with a positive q12w treatment status will be the "**efficacy/safety approach**", conducted using the FAS as described below.

The estimate for the proportion of subjects with a positive q12w treatment status will be derived from Kaplan-Meier (KM) time-to-event analyses for the event 'first q8w-need', imputing the 'q8w-need' in case of missing or confounded data attributable to lack of efficacy and/or lack of safety.

The proportion of subjects with a positive q12w treatment status will be derived as follows requiring 'sufficient duration of effect' (as assessed by q8w-need) together with 'sufficient efficacy and safety':

- For the 'sufficient duration of effect' requirement subjects will need to have the status of 'q8w need =no' at Weeks 32, 36 and 48 unless the 'q8w need = yes' is confounded by reasons other than lack of efficacy and/or safety (see censoring details below)
- The requirement regarding 'sufficient efficacy and safety' will be addressed by considering subjects – even without an explicit 'q8w-need = yes' – as having a negative q12w status in case any of the following confounding factors is attributable to lack of efficacy and/or lack of safety of the study treatment: early treatment/study discontinuation, missed DAA. The q8w-need assessment is imputed as "Yes" at the DAA visit following early treatment/study discontinuation due lack of efficacy and/or lack of safety of the study treatment (applies to both missing and non-missing DAAs).

Intercurrent events associated with missing or confounded data regarding the q12w treatment status that are attributable to reasons other than lack of efficacy and/or safety are described below, together with the corresponding data handling strategies:

- Early treatment/study discontinuation: censoring at the last valid DAA visit while on treatment/study

- Single missed visit with a relevant DAA: censoring at the last valid DAA prior to the missed visit
- Prohibited concomitant medications/procedures: censoring at the last valid DAA prior to the corresponding application
- Discrepancy between DAA by investigator and the actual treatment received: censoring at the corresponding visit
- Other treatment allocations/applications deviating from the concept of 'treatment allocation according to disease activity': censoring at the last valid DAA at or prior to the deviating visit.

Censoring rules related to the q12w treatment status analysis are described in Section 5.6.

The proportion of subjects with a positive q12w treatment status at Week 52 will be presented together with two-sided 95% confidence intervals (see Section 5.4.2.3).

The outcome of the Kaplan-Meier analysis will be presented graphically by the estimated q12w-probability over time, i.e. at each DAA visit.

While for the analysis of the overall q12w proportion all subjects in the FAS will be considered, the analysis of the proportion of subjects maintained at q12w up to Week 52, within those subjects that qualified for q12w at Week 36, is based on the subset of FAS subjects with no identified q8w-need at Week 32 and Week 36. For this subset of subjects a valid Week 36 DAA is required, while missing the Week 32 assessment is considered as no q8w-need.

2.6.3 Handling of missing values/censoring/discontinuations

The details regarding handling of missing values and discontinuations, including the timing of censoring within the time-to-event analyses for the event 'first q8w-need', are specified in the previous section.

Remark: Subjects without any valid DAA are considered censored at baseline for the overall q12w proportion and for the analysis of the predictive value of the first q12w cycle.

From an estimand perspective, the impact of failing study completion according to the protocol due to lack of efficacy/safety is considered adequately reflected by a negative q12w-status.

2.6.4 Supportive analyses

Considerations around the occurrence of potential confounding effects impacting the assessment of the q12w treatment status at Week 52 lead to the development of different estimating approaches. Supportive analyses will be performed on the FAS using alternative methods of handling missing or confounded data:

- **'Efficacy only' approach**: approach with 'q8w-need' imputation only in case the reason for a missing or confounded q12w status is attributable to lack of efficacy of the study treatment. In case of a corresponding safety reason the subject is censored at the last valid DAA.
- **'As observed' approach:** analysis without 'q8w-need' imputation.

Additionally, analyses described in Section 2.6.2 conducted using the FAS will be repeated using the PPS to assess the consistency of the assessment of the q12w proportions when looking only at those subjects who adhere to the protocol.

Subgroup analyses will be conducted as well to assess the consistency of the assessment of the q12w proportions across various subgroups described in Section 2.2.1, considering the FAS only and the efficacy/safety approach only. It will include assessment of the impact of COVID-19 pandemic with the subgroup analyses of exposed/non-exposed subjects and impacted/non-impacted subjects as defined in Section 2.5.4.3.

In addition, potential confounding effects related to COVID-19 impacting the assessment of the q12w treatment status at Week 52 led to the development of a supplementary estimand. The analysis will be repeated using the FAS and the efficacy/safety approach only, after excluding subjects whose DAA assessment has been impacted by COVID-19 (e.g., DAA not performed due to COVID-19 site impact).

2.7 Analysis of secondary efficacy endpoints

2.7.1 Secondary efficacy endpoints

Secondary efficacy endpoints related to BCVA, dosing regimen, anatomy or status of diabetic retinopathy are listed below. All endpoints that consider data from post-Week 52 visits will not be part of the primary analysis at Week 52. These endpoints will only be analyzed in the year-2 CSR (Week 100).

[redacted]

Secondary efficacy endpoints based on BCVA:

- Change from baseline in BCVA at each visit up to Week 100.
- Average change from baseline in BCVA over the period Week 88 to Week 100 (year-2 analysis only). For each subject this endpoint is derived as the average of the changes from baseline to Weeks 88, 92, 96, 100.
- Average change from baseline in BCVA over the period Week 4 to Week 52/100. For each subject this endpoint is derived as the average of the changes from baseline to each post-baseline visit between Week 4 and Week 52/100.
- Average change from baseline in BCVA over the period Week 20 to Week 52/100 and Week 28 to Week 52/100. For each subject those endpoints are derived as the average of the changes from baseline to each post-baseline visit between Week 20 and Week 52/100, and between Week 28 and Week 52/100.
- Number and percentage of subjects with a gain in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline to each post-baseline visit

  Note: Subjects with BCVA value of 84 letters or more at a post-baseline visit will be considered as responders for the corresponding endpoint. This is to account for a ceiling effect, e.g. for the' ≥15-letter gain' endpoint, for those subjects with BCVA values at baseline ≥70 letters.

- Time to achieve gain in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline (or reaching a score of 84 or more)
- Number and percentage of subjects with a loss in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline to each post-baseline visit
- Number and percentage of subjects with an absolute BCVA ≥73 ETDRS letters at each post-baseline visit

Secondary efficacy endpoints related to dosing regimen:

- Proportion of subjects maintained at q12w up to Week 64 (after three q12w treatment intervals) and 100 (for brolucizumab treatment arms only). This endpoint will only be assessed in the year-2 analysis (Week 100), using the KM method as described for the key secondary endpoint at Week 52 (Section 2.6.2)
- Proportion of subjects maintained at q12w up to Week 64 (after three q12w treatment intervals), within those subjects that qualified for q12w at Week 36 (for brolucizumab treatment arms only). This endpoint will only be assessed in the year-2 analysis (Week 100), using the KM method as described for the key secondary endpoint at Week 52 (Section 2.6.2)
- Number and percent of subjects with q8w treatment need status assessed at Week 32
- Treatment status at Week 100. This endpoint will only be assessed in the year-2 analysis (Week 100)

Secondary efficacy endpoints related to anatomy:

- Change from baseline in Central Subfield Thickness (CSFT, as determined by SD-OCT from the central reading center) at each assessment visit
- Average change from baseline in CSFT over the period Week 40 through Week 52. For each subject this endpoint is derived as the average of the changes from baseline to Weeks 40, 44, 48, 52. Then the same will be derived over the period Week 88 through Week 100, considering the average of the changes from baseline to Weeks 88, 92, 96, 100. This endpoint will only be assessed in the year-2 analysis (Week 100)
- Average change from baseline in CSFT over the period Week 4 to Week 52/100
- Proportion of subjects with normal CSFT thickness (<280 microns) at each assessment visit
- Proportion of subjects with presence of SRF, IRF and simultaneous absence of SRF and IRF (i.e., presence of SRF and/or IRF) at each assessment visit
- Proportion of subjects with presence of leakage on FA at Weeks 52 and 100

Secondary efficacy endpoints related to the status of Diabetic Retinopathy (see Section 2.7.2.1):

- Proportion of subjects with a ≥2- and ≥3-step improvement or worsening from baseline in the ETDRS Diabetic Retinopathy Severity Scale (DRSS) score at each assessment visit

- Incidence of progression to proliferative diabetic retinopathy (PDR) as assessed by ETDRS-DRSS score ≥61 by Week 52 and Week 100, among non-PDR subjects at screening

2.7.2 Statistical hypothesis, model, and method of analysis

2.7.2.1 Confirmatory testing related to additional secondary efficacy endpoints

Confirmatory hypothesis testing for additional secondary endpoints will be performed in case the proof of non-inferiority related to BCVA is successful for the four hypotheses specified above (Section 2.5.2) for the primary and first key secondary endpoints (corresponding to H1, H2, H3 and H4 in Figure 2-1).

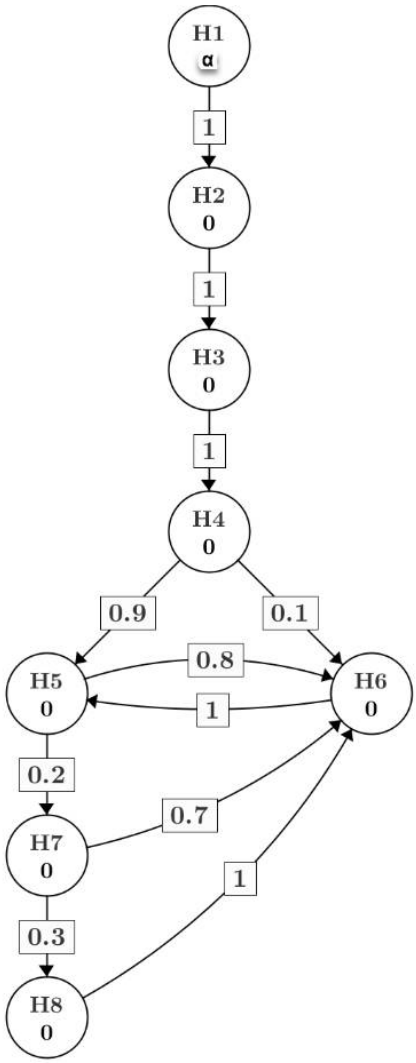
The additional hypotheses are linked to the endpoints below:

- H5. Average change from baseline in CSFT over the period Week 40 through Week 52 in the study eye
- H6. Absence of Fluid in the study eye at Week 52 (no= absence of SRF and IRF)
- H7. Change from baseline in CSFT at Week 4 in the study eye
- H8. Average change from baseline in BCVA over the period Week 40 through Week 52 in the study eye

All tests will be one-sided tests for superiority of brolucizumab 6 mg vs aflibercept 2 mg only (not brolucizumab 3 mg vs aflibercept 2 mg), i.e.,

- for greater reductions in the CSFT change from baseline (ANOVA as specified in Section 2.7.2.2) with brolucizumab compared to aflibercept
- for a higher proportion of subjects with absence of fluid (logistic regression as specified in Section 2.7.2.2) with brolucizumab compared to aflibercept
- for greater gains in the BCVA change from baseline (ANOVA as specified in Section 2.7.2.2) with brolucizumab compared to aflibercept

Figure 2-1 Multiple testing strategy



- Hypotheses H_1,..., H_8 are represented by circles with the initial significance levels. The arrow represents the direction in which the significance level is propagated throughout the graph and the number in the square box represents the proportion of the propagated significance level.

All the tests are performed at the level resulting from the graphical procedure. If a tested null hypothesis is rejected at the local significance level assigned to this null hypotheses, the alpha is passed on to other null hypotheses as per the graph.

As described in Section 2.5.2, the first four hypotheses will be tested sequentially in the order of their numbering (H_n, n=1, 2, 3, 4), i.e., confirmatory testing of the second, third or fourth hypotheses requires rejection of each preceding null hypothesis.

If each of the first four null hypotheses is rejected at a one-sided significance level of 0.025, the entire alpha will be distributed between the null hypotheses related to the superiority testing of H_5 (90% of 0.025 = 0.0225), and H_6 (10% of 0.025 = 0.0025). This split is chosen by balancing out prior expectations about the study outcomes and the clinical importance of the endpoints.

The family-wise type I error rate will be controlled at the one-sided 2.5% level across the tested null hypotheses using the closed testing procedure specified by Figure 2-1 using the graphical method of Bretz, et al. (Bretz, et al 2009).

The basis for these tests for superiority will be the FAS with LOCF imputation/replacement of missing/censored data. For subjects who discontinue study treatment but continued in the study, data collected after the switch to alternative DME treatment in the study eye will be censored for the primary analysis.

2.7.2.2 General analysis specifications for secondary efficacy endpoints

All secondary efficacy endpoints listed in the above Section 2.7.1 will be summarized and presented descriptively, based on the FAS with LOCF imputation for missing or censored data if not otherwise specified. Details on data handling, such as missing values, are described in Section 2.7.3.

The impact of COVID-19 pandemic on CSFT will be assessed with the subgroup analyses of exposed/non-exposed subjects and impacted/non-impacted subjects as defined in Section 2.5.4.3.

Continuous endpoints:

The continuous secondary endpoints related to BCVA and CSFT will be analyzed using ANOVA models. The estimates of least square means for each treatment and for the treatment differences brolucizumab – aflibercept, including 95% CIs for the treatment differences, will be presented.

For the ANOVA analysis of BCVA-related endpoints, baseline BCVA (≤65, >65 letters) and age category (<65, ≥65 years) will be considered as class variables. For the ANOVA analysis of CSFT, baseline CSFT (<450, ≥450-<650, ≥650 µm) will be used instead of baseline BCVA as a class variable.

The line plot on LSM (± SE) by visit will also be provided for all treatment arms.

Categorical variables:

For binary endpoints, frequency tables (count and percentage) will be provided by time point. In addition, proportions and treatment differences in proportions along with 95% CIs will be presented for each time point using a logistic regression with treatment, the corresponding baseline status (similar to the ones specified for the ANOVA models) and age categories as fixed effects.

Bar chart will be plotted by visit and treatment arm.

Time-to event variables:

Time-to-event variables such as the time to achieve gain in BCVA of ≥5 (respectively ≥10 and ≥15) letters from baseline (or reaching a score of 84 or more) will be analyzed using KM analysis. KM estimates on percent of subjects who achieve gain, together with 95% CI will be presented by visit. The median time (95% CI) to gain will also be constructed by treatment arm. KM curves presenting the cumulative probability of subjects with gain of ≥5 (respectively ≥10 and ≥15) letters from baseline will be provided by treatment arm.

2.7.2.3 ETDRS DRSS Score

Definition of Endpoints

The following endpoints related to diabetic retinopathy (DR) status will be analyzed:

- Subject status regarding a ≥2- and ≥3-step improvement or worsening from baseline in the ETDRS DRSS score at each assessment visit
- Incidence of progression to PDR as assessed by ETDRS-DRSS score of at least 61 by Week 52 and Week 100 (among non-PDR subjects at screening)

Those endpoints will be derived from the ETDRS-DRSS score assessed by the central reading center based on colour fundus (CF) photography images in the study eye at screening, Weeks 28, 52, 76 and exit/premature discontinuation visit.

When the ETDRS-DR severities are evaluable, they will be categorized using the following scores:

Table 2-3 Definition of DRSS: original scale

DRSS scale	Definition
10	DR absent
20	Microaneurysms only
35	Mild non-proliferative diabetic retinopathy (NPDR)
43	Moderate NPDR
47	Moderately severe NPDR
53	Severe NPDR
61	Mild PDR
65	Moderate PDR
71	High-Risk PDR
75	Very high risk PDR
81	Advanced PDR
85	Very advanced PDR

Other recorded DRSS values (code 98: Indeterminable due to missing images, 99: Indeterminable due to upgradable images, 00: No images received) that are not related to an evaluable DR severity level will be handled as missing.

All DRSS values will be converted into a 12-point scale as defined in Table 2-4.

Table 2-4 Definition of DRSS: 12-point scale

12-point scale	Definition	Original DRSS
1	DR absent	10
2	Microaneurysms only	20
3	Mild NPDR	35
4	Moderate NPDR	43
5	Moderately severe NPDR	47

12-point scale	Definition	Original DRSS
6	Severe NPDR	53
7	Mild PDR	61
8	Moderate PDR	65
9	High-Risk PDR	71
10	Very high-Risk PDR	75
11	Advanced PDR	81
12	Very advanced PDR	85

DR= diabetic retinopathy, DRSS= diabetic retinopathy severity score, NPDR= non-proliferative diabetic retinopathy, PDR= proliferative diabetic retinopathy.

Table 2-5 and Table 2-6 describe the definition of a 2-step and a 3-step change, respectively, for each (non-missing) baseline and post-baseline ETDRS based on the 12-point scale, as defined below:

- ≥2-step improvement: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≤-2
- ≥3-step improvement: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≤-3
- ≥2-step worsening: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≥2
- ≥3-step worsening: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≥3

Table 2-5 Definition of 2-step change in DRSS on the 12-point scale

	Post-baseline		
Baseline	≥2-step improvement	No change or change <2 steps	≥2-step worsening
1	-	1, 2	3 or higher
2	-	1, 2 or 3	4 or higher
3	1	2, 3, or 4	5 or higher
4	1 or 2	3, 4, or 5	6 or higher
5	3 or lower	4, 5, or 6	7 or higher
6	4 or lower	5, 6, or 7	8 or higher
7	5 or lower	6, 7, or 8	9 or higher
8	6 or lower	7, 8, or 9	10 or higher
9	7 or lower	8, 9, or 10	11 or 12
10	8 or lower	9, 10, or 11	12
11	9 or lower	10, 11, or 12	-
12	10 or lower	11, 12	-

Table 2-6 Definition of 3-step change in DRSS on the 12-point scale

Baseline	Post-baseline		
	≥3-step improvement	No change or change <3 steps	≥3-step worsening
1	-	1, 2 or 3	4 or higher
2	-	1, 2, 3 or 4	5 or higher
3	-	1, 2, 3, 4 or 5	6 or higher
4	1	2, 3, 4, 5 or 6	7 or higher
5	1 or 2	3, 4, 5, 6 or 7	8 or higher
6	3 or lower	4, 5, 6, 7 or 8	9 or higher
7	4 or lower	5, 6, 7, 8 or 9	10 or higher
8	5 or lower	6, 7, 8, 9 or 10	11 or 12
9	6 or lower	7, 8, 9, 10 or 11	12
10	7 or lower	8 or higher	-
11	8 or lower	9 or higher	-
12	9 or lower	10 or higher	-

Analysis method

All DRSS analyses will be based on the 12-point scale shown in Table 2-4.

Proportions of subjects with ≥2- and ≥3-step improvement or worsening from baseline will be summarized using the FAS by assessment visit. Bar chart will be plotted by assessment visit and treatment arm.

For the proportions of subjects with ≥2-step change from baseline at Week 52 (and similarly for ≥3-step change), the 95% confidence intervals (CIs) for the proportions in all treatment arms, the differences in proportions between brolucizumab and aflibercept treatment arms and the 95% CI for the difference will be calculated using a logistic regression with treatment, the corresponding baseline DRSS (score ≤43, ≥47 from the original scale or ≤4, ≥5 from the 12-level scale) and age category (<65, ≥65 years) as fixed effects.

The impact of COVID-19 pandemic on the proportion of subjects with ≥2-step improvement or worsening from baseline will be assessed with the subgroup analyses of exposed/non-exposed subjects and impacted/non-impacted subjects as defined in Section 2.5.4.3.

The proportion of subjects who progress to PDR, as assessed by ETDRS-DRSS score ≥61 by Week 52 and Week 100, will be summarized among the subset of non-PDR subjects at screening (ETDRS-DRSS score <61 at screening).

2.7.3 Handling of missing values/censoring/discontinuations

Missing data for all the secondary efficacy endpoints will be imputed using the LOCF method unless specified otherwise.

For the LOCF method, missing data will be imputed by the value of the last available non-missing post-baseline observation. For subjects who discontinue treatment but continue in the study, data collected after the start of alternative DME treatment in the study eye will be

censored for the analysis. Censored data will be replaced by the last available observation prior to the start of alternative DME treatment in the study eye.

Missing baseline values will not be imputed. For subjects with no post-baseline values (scheduled or unscheduled), the baseline value will be carried forward, as a conservative approach.

For endpoints related to presence of SRF and/or IRF, if baseline visit is reported as ”Cannot Grade”, then it will be considered as ”Absent”; if post-baseline visit is reported as ”Cannot Grade”, then it will be considered as missing and LOCF method for imputation will be applied.

For the presence of leakage on FA, if baseline visit is reported as ”Cannot Grade”, then it will be considered as missing; if post-baseline visit is reported as ”Cannot Grade”, then it will be considered as missing and LOCF method for imputation will be applied.

2.8 Safety analyses

Safety endpoints are based on the variables from safety assessments, which include:

- Extent of exposure (see Section 2.4.1)
- Adverse events
- Ophthalmic examinations
- Vital signs
- Laboratory results
- Imaging parameters

There are no formal safety hypotheses in this study. All safety analyses will be performed using the Safety Analysis Set.

Except for imputation of partial dates for AEs, no imputations will be performed for missing values in the safety analyses.

In all summary tables, unless otherwise specified (e.g. for AE tables), data collected after the subject discontinued study treatment and started alternative DME treatment in the study eye will be censored (data on the day the subject started alternative DME treatment will be included).

2.8.1 Adverse events (AEs)

A treatment-emergent adverse event (TEAE) is defined as any adverse event that develops after exposure to the study treatment or any event already present that worsens following exposure to the study treatment. Only treatment-emergent adverse events will be presented in the summary tables.

Adverse events will be coded using the MedDRA dictionary and presented by system organ class (SOC), preferred term (PT) and treatment arm. Treatment-emergent AEs will be analyzed based on the number and percentage of subjects with at least one AE in the category of interest.

The number (and proportion) of subjects with TEAEs will be summarized at each analysis timepoint (Week 52, Week 100) in the following ways:

Table 2-7 TEAE summary

TEAE summary	AE categories		
	Ocular AE in the study eye	Ocular AE in the fellow eye	Non-ocular AE
AEs by primary SOC and PT	Y#		Y#
AEs by primary SOC and PT (including events with onset date after start of alternative DME treatment)	Y	Y	Y
Frequent AEs by PT†	Y		Y
AEs by maximum severity, SOC and PT	Y		Y
AEs related to study treatment by SOC and PT	Y		Y
AEs related to injection procedure by SOC and PT	Y		
AEs leading to permanent discontinuation of study treatment by SOC and PT	Y		Y
AEs leading to temporary interruption of study treatment by SOC and PT	Y		Y
SAEs by SOC and PT	Y#		Y#
SAEs by SOC and PT (including events with onset date after start of alternative DME treatment)	Y	Y	Y
SAEs related to study treatment by SOC and PT	Y		Y
SAEs related to injection procedure by SOC and PT	Y		

†≥2 % (or other cutting point as appropriate) in any treatment group for a given PT.

including separate summary tables for exposed/non-exposed and impacted/non-impacted subjects to COVID-19 as defined in Section 2.5.4.3

In all summary tables listed above, unless otherwise specified, data collected after the subject discontinued study treatment and started alternative DME treatment in the study eye will be censored.

If an AE started on the same day as the start of alternative DME treatment for a subject, the AE will be excluded from the summary table, unless this AE led to study drug withdrawal (in such a case, the AE would be included in the analysis).

Subject listings of all adverse events will be provided. Deaths and SAEs (i.e., other serious or clinically significant non-fatal adverse events) will be listed separately.

The MedDRA version used for reporting the AEs will be described in a footnote.

2.8.1.1 Adverse events of special interest / grouping of AEs

Incidence of adverse events of special interest (AESI) will be tabulated by treatment arm.

AESIs will be identified via the RTH258 electronic case retrieval strategy (eCRS). The eCRS that is current at the time the database lock will be used and AESIs will be identified where the flag Core Safety Topic Risk (SP) = 'Y'.

2.8.1.2 Adverse event reporting for clinical trial safety disclosure

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on TEAEs which are not serious adverse events with an incidence greater than 5% and on TEAEs and SAEs suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population.

If for the same subject, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is >1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE (respectively non-SAE) has to be checked in a block e.g., among AEs in a ≤1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment, and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.8.2 Deaths

A summary of treatment emergent deaths will be presented by primary SOC and PT.

All deaths recorded in the clinical database will be listed.

2.8.3 Laboratory data

Laboratory data will be presented graphically using boxplots of absolute change from baseline values by treatment arm and visit. No summary by visit tables will be provided.

A summary table with counts and percentage of subjects satisfying the criteria representing clinically relevant abnormalities given in Section 5.3 at any visit will be presented. A listing for subjects satisfying at least one criterion in Table 5-1 at any visit will also be presented.

2.8.4 Other safety data

2.8.4.1 Ophthalmic examinations

Descriptive summaries of pre-injection change from baseline in intraocular pressure (IOP) values for the study eye will be presented graphically at each study visit by treatment arm, considering line plots of the mean change in IOP values with error bars representing ± SE. The x-axis will be study visit and the y-axis will be the change from Baseline value. No summary by visit tables will be provided.

The number and percentage of subjects with pre-injection IOP >30 mmHg at any visit will be summarized.

Post-injection IOP is to be assessed approximately within 60 minutes after injection and if ≥25 mmHg, the assessment should be repeated until back to normal. Summary tables with

counts and percentage of subjects with an IOP increase of ≥10, ≥20 mmHg from pre-injection to post-injection at any visit for the study eye will be presented.

A summary table with counts and percentage of subjects with observed pre-injection IOP ≥ 21 mmHg at 3 consecutive scheduled visits will be presented.

A visit with missing pre-injection IOP is considered to meet the ≥21 mmHg criterion if the preceeding and the following visits meet the criterion that pre-injection IOP ≥21 mmHg. For example, if schedule visit X has missing pre-injection IOP and pre-injection IOP ≥21 mmHg is observed for both visit X-1 and X+1, the subject is considered to meet the criteria at visit X as well.

A listing for subjects with any post-injection IOP increase of ≥10 mmHg from pre-injection IOP and a listing of subjects with any IOP >30 mmHg will be presented.

The abnormal findings via slit-lamp and indirect fundus examinations deemed as clinically significant by the investigator and reported as AE/SAE will be included in the safety analysis on AE/SAE.

2.8.4.2 Loss in BCVA

The number and percentage of subjects with a loss in BCVA ≥15, ≥30 letters (study eye) from baseline to each visit, to the last visit, and maximum loss at any visit will be presented.

BCVA data (study eye) for subjects presenting loss in BCVA ≥15 letters from baseline at any post-baseline visit will be listed.

2.8.4.3 Vital signs

A summary table with counts and percentage of subjects satisfying the criteria given in Table 5-2 of the Section 5.3 at least one visit will be presented. A listing for subjects satisfying at least one criterion in Table 5-2 will also be presented.

A line plot of mean change from baseline in the vital sign parameter by study visit and treatment arm with error bars representing ±1 standard error will be presented. The x-axis will be study visit and the y-axis will be the mean change from baseline value.

2.8.4.4 Imaging parameters

Imaging parameters in the study eye typically associated with intraocular inflammation and/or retinal vascular occlusion as assessed by the CRC will be listed per visit for subjects with AESI in the study eye only. No summary table will be provided.

2.9 Pharmacokinetic endpoints

Not Applicable.

2.10 Anti-drug antibodies

Collection of blood sample for ADA assessment for brolucizumab will be done at Screening, Weeks 4, 12, 24, 36, 52, and 76 prior to the injection/sham, and at exit/premature discontinuation.

Systemic exposure of brolucizumab will be measured concomitantly with ADA levels for interpretation purposes, no pharmacokinetic parameters will be determined from brolucizumab systemic exposure. Systemic exposure data will be summarized and listed.

2.11 Patient-reported outcomes

The Visual Function Questionnaires (VFQ-25) will be scored (total and subscale scores) at Baseline and Weeks 28, 52, 76 and 100 visits. Absolute scores and the absolute changes from baseline will be calculated and summarized descriptively using the FAS.

Further details on the scoring algorithm and analysis are provided below.

Each subscale score has a range of 0 to 100 inclusive and will be calculated from the re-scaled raw data as described in Table 2-8. A missing response will not be re-scaled (except for the response to question 15c, see below, which will be re-set to 0 if the response to question 15b is 1).

The answers to questions will be re-scaled as follows to calculate the total and subscale scores.

Table 2-8 Rescaling of VFQ-25 questions

Answer to question	Rescaling for questions 1, 3, 4 and 15c	Rescaling for question 2	Rescaling for questions 5-14, 16 and 16a	Rescaling for questions 17-25
1	100	100	100	0
2	75	80	75	25
3	50	60	50	50
4	25	40	25	75
5	0	20	0	100
6	N/A	0	N/A*	N/A

Note: * 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". Subscales will be calculated where at least one of the (re-scaled) questions contributing to that subscale is non-missing, and otherwise set to missing.

- Note that the answer to question 15c will subsequently be adjusted based on the answer to question 15b.
    - If the answer to 15b is 1 then the answer to 15c will be re-set to 0.
    - If the answer to 15b is 2 or 3 then the answer to 15c will be re-set to missing

The general health rating is the re-scaled answer to question 1.

The scales and corresponding questions are shown in Table 2-9.

Table 2-9 Questions contributing to VFQ subscales

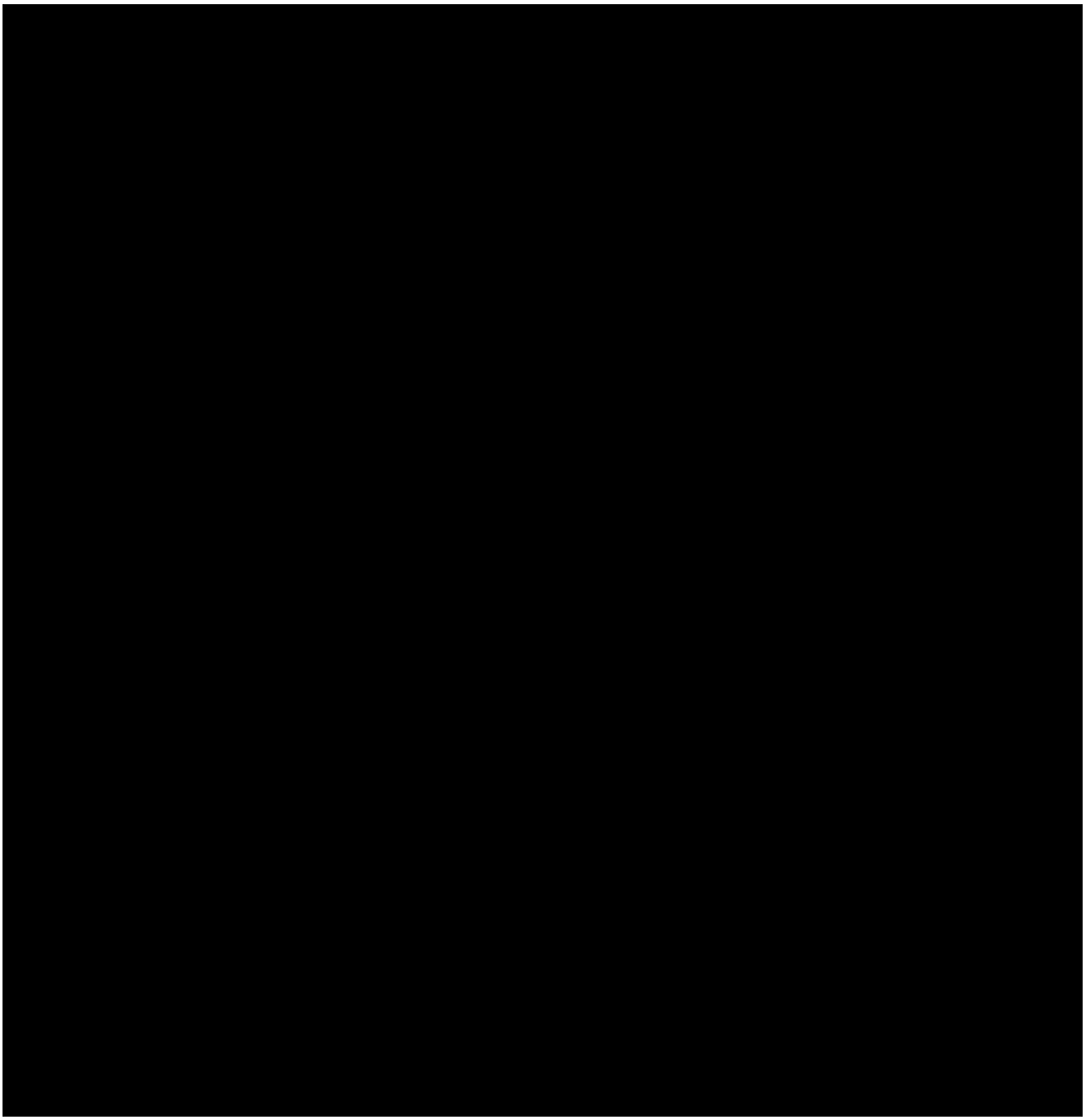
Subscale	Questions
General vision	2
Ocular pain	4 and 19
Near activities	5, 6 and 7
Distance activities	8, 9 and 14
Social functioning	11 and 13
Mental health	3, 21, 22 and 25
Role difficulties	17 and 18
Dependency	20, 23 and 24
Driving	15c, 16 and 16a
Color vision	12
Peripheral vision	10

The composite score is the average of the 11 subscales shown in Table 2-9. It will be set to missing if at least six of the subscales are missing.

Descriptive summary statistics for change from baseline to post baseline VFQ assessments will be presented using the FAS for the composite and subscale scores. Mean changes from baseline to each post baseline VFQ assessments visits will be compared between the brolucizumab arms and the aflibercept arm. Appropriate statistical methods (e.g. pairwise ANCOVA model with treatment as a fixed effect factor and corresponding baseline value of the endpoint in the model)

will be used for treatment arm comparisons. Additionally, descriptive statistics will also be presented for the general health score. All analyses will be performed on the subscales values.

The VFQ-25 composite score and subscale scores will not be listed.



2.13 Interim analysis

The analysis based on the Week 52 data will be the primary efficacy and safety analysis for this study. The database includes all data up to Week 52 from when all randomized subjects have completed the Week 52 visit or terminated the study prior to (or at) Week 52.

The results of this analysis will be reported in theCSR and will include the analysis of the overall population and subgroup analyses Japan vs non-Japan for the following data:

- Subject disposition (Section 2.3.1)
- Demographics and baseline characteristics (Section 2.3.2)
- Study treatment exposure (Section 2.4.1)
- Subgroup analysis of the primary and key secondary efficacy variables (Section 2.2.1)
- Change from baseline in BCVA at each visit up to Week 52 (Section 2.7)
- Change from baseline in CSFT at each visit up to Week 52 (Section 2.7)
- Proportion of subjects with presence of SRF, IRF and simultaneous absence of SRF and IRF (i.e., presence of SRF and/or IRF) at each assessment visit (Section 2.7)
- Proportion of subjects with a ≥2- and ≥3-step improvement or worsening from baseline in the DRSS score at each assessment visit (Section 2.7)
- The number (and proportion) of subjects with ocular AEs/SAEs in the study eye and non-ocular AEs/SAEs, up to Week 52 (Section 2.8.1)

A second planned interim analysis may be performed by locking the Week 76 data in case of regulatory request of supplemental data to be submitted during the review period.

3 Sample size calculation

A sample size of 160 subjects per arm will allow demonstration of non-inferiority (NIM of 4 ETDRS letters) of brolucizumab 6 mg or 3 mg vs. aflibercept 2 mg with respect to the BCVA change from baseline at Week 52, with 90% power (disregarding the dependence within the sequential testing procedure, i.e. local power for 3 mg) at a one-sided alpha level of 0.025, assuming equal means and a common standard deviation of 11 letters. Assuming that averaging over the 4 time points will not lead to an increase in the standard deviation a power of at least 90% can also be expected for its corresponding non-inferiority claim.

To account for a drop-out rate of 10%, a total of 534 (178 per arm) subjects will need to be randomized.

4 Change to protocol specified analyses

There is no change to the protocol specified analyses in terms of endpoints.

Confirmatory hypothesis testing in relation to additional secondary endpoints is introduced in Section 2.7.2.1. Some changes compared to the protocol specified analyses are considered in the current statistical analysis plan before database lock in order to implement the Novartis internal process on SAP simplification (LEAN):

Protocol section	Protocol wording	Change in the SAP
9.2	Demographics and baseline characteristics will be summarized with descriptive statistics for all analysis sets by treatment group and overall.	Demographics and baseline characteristics will be summarized with descriptive statistics for the FAS by treatment arm and overall.
9.3	Descriptive statistics for exposure to study treatment will be provided for the safety set, FAS and PPS	Descriptive statistics for exposure to study treatment will be provided for the Safety set.
[redacted]		
9.5.2	Laboratory data and vital signs will be summarized by presenting shift tables using extended normal ranges (as provided by the central laboratory) with thresholds representing clinical relevant abnormality and by presenting descriptive statistics of raw data and change from baseline. Values outside the extended normal range will be listed by subject and treatment arm and flagged in data listings.	No summary by visit tables will be provided. A summary table with counts and percentage of subjects satisfying the criteria representing clinically relevant abnormalities at any visit will be presented. A listing for subjects satisfying at least one criterion in Table 5-1 at any visit will also be presented.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

No imputation will be made to the start date and end date of study treatment.

5.1.2 AE date imputation

5.1.2.1 AE start date imputation

The following table explains the notation used in the logic matrix below. Please note that completely missing start dates will not be imputed.

	Day	Month	Year
Partial Adverse Event Start Date	Not used	MON	YYYY
Treatment Start Date	Not used	TRTM	TRTY

The following matrix explains the logic behind the imputation.

	MON MISSING	MON < TRTM	MON = TRTM	MON > TRTM
YYYY MISSING	(1) No convention	(1) No convention	(1) No convention	(1) No convention

	MON MISSING	MON < TRTM	MON = TRTM	MON > TRTM
YYYY < TRTY	(2.a) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start
YYYY = TRTY	(4.a) Uncertain	(4.b) Before Treatment Start	(4.c) Uncertain	(4.c) After Treatment Start
YYYY > TRTY	(3.a) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start

Before imputing AE start date, find the AE start reference date.

1. If the (imputed) AE end date is complete and the (imputed) AE end date < treatment start date then AE start reference date = min (informed consent date, earliest visit date).
2. Else AE start reference date = treatment start date.

Impute AE start date -

1. If the AE start date year value is missing, the date uncertainty is too high to impute a rational date. Therefore, if the AE year value is missing, the imputed AE start date is set to to NULL.
2. If the AE start date year value is less than the treatment start date year value, the AE started before treatment. Therefore:
   a. If AE month is missing, the imputed AE start date is set to the mid-year point (01JulYYYY).
   b. Else if AE month is not missing, the imputed AE start date is set to the mid-month point (15MONYYYY).
3. If the AE start date year value is greater than the treatment start date year value, the AE started after treatment. Therefore:
   a. If the AE month is missing, the imputed AE start date is set to the year start point (01JanYYYY).
   b. Else if the AE month is not missing, the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).
4. If the AE start date year value is equal to the treatment start date year value:
   a. And the AE month is missing the imputed AE start date is set to the AE reference start date + 1 day.
   b. Else if the AE month is less than the treatment start month, the imputed AE start date is set to the mid-month point (15MONYYYY).
   c. Else if the AE month is equal to the treatment start date month or greater than the treatment start date month, the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).

If complete (imputed) AE end date is available and the imputed AE start date is greater than the (imputed) AE end date, then imputed AE start date should be set to the (imputed) AE end date.

5.1.2.2 AE end date imputation

1. If the AE end date month is missing, the imputed end date should be set to the earliest of the (31DECYYYY, date of death).

2. If the AE end date day is missing, the imputed end date should be set to the earliest of the (last day of the month, date of death).
3. If AE year is missing or AE is ongoing, the end date will not be imputed.
4. If the imputed AE end date is less than the existing AE start date then use AE start date as AE end date.

5.1.3 Concomitant medication date imputation

5.1.3.1 Concomitant medication start date

In order to classify a medication as prior and prior/concomitant, it may be necessary to impute the start date.

Completely missing start dates will be set to one day prior to treatment start date. As a conservative approach, such treatments will be classified as prior and concomitant (and summarized for each output).

Concomitant treatments with partial start dates will have the date or dates imputed.

The following table explains the notation used in the logic matrix

	Day	**Month**	**Year**
Partial CMD Start Date	Not used	MON	YYYY
Treatment Start Date	Not used	TRTM	TRTY

The following matrix explains the logic behind the imputation.

	MON MISSING	**MON < TRTM**	**MON = TRTM**	**MON > TRTM**
YYYY MISSING	(1) Uncertain	(1) Uncertain	(1) Uncertain	(1) Uncertain
YYYY < TRTY	(2.a) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start
YYYY = TRTY	(4.a) Uncertain	(4.b) Before Treatment Start	(4.a) Uncertain	(4.c) After Treatment Start
YYYY > TRTY	(3.a) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start

1. If the CM start date year value is missing, the imputed CM start date is set to one day prior to treatment start date.
2. If the CM start date year value is less than the treatment start date year value, the CM started before treatment. Therefore:
   a. If the CM month is missing, the imputed CM start date is set to the mid-year point (01JulYYYY).
   b. Else if the CM month is not missing, the imputed CM start date is set to the mid-month point (15MONYYYY).
3. If the CM start date year value is greater than the treatment start date year value, the CM started after treatment. Therefore:

a. If the CM month is missing, the imputed CM start date is set to the year start point (01JanYYYY).
b. Else if the CM month is not missing, the imputed CM start date is set to the month start point (01MONYYYY).

4. If the CM start date year value is equal to the treatment start date year value:
a. And the CM month is missing or the CM month is equal to the treatment start date month, then the imputed CM start date is set to one day prior to treatment start date.
b. Else if the CM month is less than the treatment start date month, the imputed CM start date is set to the mid-month point (15MONYYYY).
c. Else if the CM month is greater than the treatment start date month, the imputed CM start date is set to the month start point (01MONYYYY).

If complete (imputed) CM end date is available and the imputed CM start date is greater than the (imputed) CM end date, then imputed CM start date should be set to the (imputed) CM end date.

5.1.3.2 Concomitant medication (CM) end date imputation

1. If the CM end date year value is missing, the date uncertainty is too high to impute a rational date. Therefore, if the CM end year value is missing or ongoing, the imputed CM end date is set to NULL.

2. Else, if the CM end date month is missing, the imputed end date should be set to the earliest of the (treatment end date, 31DECYYYY, date of death).

3. If the CM end date day is missing, the imputed end date should be set to the earliest of the (treatment end date, last day of the month, date of death).

4. If the imputed CM end date is less than the existing CM start date, use the CM start date as the imputed CM end date.

5.1.4 Medical history date of diagnosis imputation

Completely missing dates and partially missing end dates will not be imputed. Partial dates of diagnosis will be compared to the treatment start date.

1. If DIAG year < treatment start date year
    a. and DIAG month is missing, the imputed DIAG date is set to the mid-year point (01JULYYYY).
2. else if DIAG month is not missing, the imputed DIAG date is set to the mid-month point (15MONYYYY).
3. If DIAG year = treatment start date year
    a. and (DIAG month is missing OR DIAG month is equal to treatment start month), the imputed DIAG date is set to one day before treatment start date.
    b. else if DIAG month < treatment start month, the imputed DIAG date is set to the midmonth point (15MON YYYY).

c. else if DIAG month > treatment start month => data error.

4. If DIAG year > treatment start date year => data error.

5.2 AEs coding/severity

AEs are coded using the MedDRA terminology.

AEs severity is assessed by investigators according to the following:

- mild: usually transient in nature and generally not interfering with normal activities
- moderate: sufficiently discomforting to interfere with normal activities
- severe: prevents normal activities

5.3 Laboratory parameters and vital signs derivations

Table 5-1 Clinically notable laboratory values

Test	Conventional Units	Critical Low	Critical High	Standard Units	Critical Low	Critical High	Non-nemeric
Calcium	mg/dL	< 6.0	> 13.0	mmol/L	< 1.50	> 3.25	
Creatinine		NA	>3xULN				
Glucose	mg/dL	< 40	> 450	mmol/L	< 2.2	> 25.0	
Potassium	mEq/L	< 2.8	> 6.2	mmol/L	< 2.8	> 6.2	
Sodium	mEq/L	< 120	> 160	mmol/L	< 120	> 160	
HCG							Negative, inconclusive
Hematocrit	%	< 20	> 60	V/V	< 0.20	> 0.60	
Hemaglobin	g/dL	< 6.0	> 20.0	g/L	< 60	> 200	
Platelet	X10E3/uL	< 50	> 999	X10E9/L	< 50	> 999	
WBC	X10E3/uL	< 2.0	> 35.0	X10E9/L	< 2.0	> 35.0	

Table 5-2 Clinically notable vital signs

Variable	Category	Critical values
Systolic blood pressure (mmHg)	High	Either >180 with an increase from baseline >30 or >200 absolute
	Low	Either <90 with a decrease from baseline >30 or <75 absolute
Diastolic blood pressure (mmHg)	High	Either >105 with an increase from baseline >20 or >115 absolute
	Low	Either <50 with a decrease from baseline > 20 or <40 absolute
Pulse rate (bpm)	High	Either >120 with an increase from baseline of >25 or > 130 absolute
	Low	Either <50 with a decrease from baseline >30 or <40 absolute

5.4 Statistical models

5.4.1 Primary and first key secondary analysis

The primary endpoint (change from baseline in BCVA at Week 52) and first key secondary endpoint (average change from baseline in BCVA over the period Week 40 through Week 52) will be analyzed using ANOVA models.

The ANOVA models will be fitted separately for brolucizumab 3mg vs. Aflibercept 2mg and for brolucizumab 6mg vs. Aflibercept 2mg.

Analysis of Variance (ANOVA)
The following ANOVA model will be used for the primary and first key secondary efficacy endpoints:
<change from Baseline in BCVA at Week 52> <average change from Baseline in BCVA from Week 40 to Week 52> = intercept + treatment + Baseline BCVA category + age category + error.
For the above analysis, the data structure is one record per subject. The SAS Proc MIXED will be used to perform the ANOVA analyses.

Mixed Model Repeated Measures (MMRM)
The following MMRM model will be used for the supportive analysis of the primary and first key secondary efficacy variables:
*<change from Baseline in BCVA at Week 52> <average change from Baseline in BCVA from Week 40 to Week 52> = intercept + treatment + Baseline BCVA category + age category + visit + treatment*visit + error.*

The SAS Proc MIXED will be used to perform the MMRM analyses.

Note: For the above MMRM analysis, the data structure is one record per FAS subject per scheduled visit. The data will include all subjects and have records for all scheduled visits, regardless of whether the assessment was missed or not at a given visit. Missing values will NOT be imputed using LOCF. Instead, the value will be passed to the model as missing.

5.4.2 Other secondary efficacy analysis

5.4.2.1 ANCOVA model for continuous variables

The continuous efficacy variables (such as VQF-25 score change from baseline) will be analyzed using an ANCOVA model adjusted for treatment, age category, and the corresponding baseline VQF-25 score.
The SAS Proc MIXED will be used to perform the ANCOVA analyses

5.4.2.2 Logistic regression for proportion variables

The binary efficacy variables will be analyzed using the logistic regression model adjusted for treatment, age category, corresponding baseline variables, and other covariates if necessary, using the FAS.

The SAS Proc LOGISTIC will be used.
Note:

- For the above analyses, the data structure is one record per subject and visit.The least square mean estimates obtained from the above model are for the log-odds ratios.
- The estimated difference in proportions and the corresponding 95% confidence intervals will be obtained by applying the bootstrap method. The pseudo SAS code to derive the treatment difference and 95% CI from the least square mean output of the fitted model will be provided in the programming specification document.

5.4.2.3 KM estimate for time to event variables

Within the brolucizumab treatment arms, the proportion of subjects maintained at q12w up to Week 52 will be estimated from Kaplan Meier time-to-event analyses for the event 'first q8w-need', applying event allocations (in case of lack of efficacy and/or lack of safety) and censoring as described in Section 2.6.2.

A corresponding 95% CI will be derived from the LOGLOG transformation, using SAS Proc Lifetest, with CONFTYPE = LOGLOG.

5.5 Rule of exclusion criteria of analysis sets

Important protocol deviations are defined in the Protocol Deviations Requirements Document. Table 5-3 includes the important protocol deviations which lead to exclusion of a subject from one or more analysis sets for the Week 52 analysis:

Table 5-3 Important protocol deviations leading to exclusion from analyses

Deviation ID	Description of Deviation	Exclusion in Analyses
M_INCL01_ICF not obtained	Written informed consent not obtained	Exclude from all analyses
P_INCL02_Age less than 18 yrs	Patient less than 18 years of age at baseline	Exclude from PP analysis
M_INCL03_Diabetes eligibility criteria	Patients without diabetes mellitus or HbA1c of more than 10% at screening or insufficient diabetes management at screening or baseline	Exclude from PP analysis
P_INCL04_No visual impairment (study eye)	Study Eye: no visual impairment due to DME as per BCVA or CSFT criteria	Exclude from PP analysis
M_EXCL01_Confounding condition in study eye	Study Eye: Confounding ocular concomitant conditions or ocular disorders	Exclude from PP analysis
M_EXCL02_Confounding concomitant medications or procedures in study eye	Study Eye: Confounding concomitant medications or procedures	Exclude from PP analysis

Deviation ID	Description of Deviation	Exclusion in Analyses
M_TRT01_Wrong IP administered	Wrong IP administered during the study	Exclude from PP analysis, unless brolucizumab 3mg was given instead of brolucizumab 6mg, or brolucizumab 6mg was given instead of brolucizumab 3mg
M_TRT02_Under-treatment during loading phase	Under-treatment during loading phase; missed active treatment (not due to any safety event)	Exclude from PP analysis
M_TRT03_Over treatment	Over treatment, received active when schedule was for sham /no treatment	Exclude from PP analysis
M_TRT04_Under-treatment after loading phase	Under-treatment after loading phase; missed active treatment (not due to any safety event)	Exclude from PP analysis if any missed active between W40 and W48 inclusive, or if at least 2 missed consecutive active doses (not due to safety); Otherwise include in all analyses
M_OTH01_Masking process not followed	Masking process not followed as per protocol with impact on data integrity	Exclude from PP analysis
M_OTH02_Any other PD	Any other protocol deviation with impact on the efficacy assessments or safety of the patient	Exclude from PP analysis
P_WITH01_Treatment but consent withdrawn	Subject withdrew consent but continued to receive study medication	Exclude from PP analysis

Table 5-4 lists the non-protocol deviations (analysis restrictions) that may lead to exclusion from per-protocol analysis. Analysis restrictions (ARs) address limitations in the evaluability which result from missing or confounded data with underlying background not qualifying as a PD (e.g. early study terminations, early treatment discontinuations, missing DAA or BCVA assessments).

Subject evaluability is based on two components:

- Exclusion from an analysis set
- Censoring of specific data points from an analysis (see Section 5.6).

The consequence of an AR on the evaluability depends on the underlying reason, while three different categories of reason are considered:

- Lack of efficacy of the study treatment (=1)
- Lack of safety / tolerability of the study treatment (=2)
- Other (=0)

Remark: Based on the concept of PD's, their underlying reason will always be '0'.

As a general rule, ARs with a reason of 1 or 2 do not lead to an exclusion from any analysis set, as a potential link between exclusion reason and treatment would constitute a source for systematic bias.

Rules of determination of ARs by programming will be specified in the Programming Data Specifications (PDS) documentation.

Table 5-4 Non-protocol deviations (analysis restrictions)

AR ID	Description of AR	Category of reason	Exclusion in Analyses
AR_EST_01	Early study termination due to lack of efficacy	1	Include in all analyses
AR_EST_02	Early study termination due to lack of safety	2	Include in all analyses
AR_EST_03	Early study termination due to reasons other than lack of efficacy/safety	0	Exclude from PP analysis if before Week 40 Otherwise include in all analyses
AR_ETD_01	Early study treatment termination due to lack of efficacy	1	Include in all analyses
AR_ETD_02	Early study treatment termination due to lack of safety	2	Include in all analyses
AR_ETD_03	Early study treatment termination due to reasons other than lack of efficacy/safety	0	Exclude from PP analysis if before Week 40 Otherwise include in all analyses
AR_MD_01	No valid BCVA assessment between Week 40 and Week 52	0	Exclude from PP analysis
AR_MD_02	Missing DAA due to lack of safety	2	Include in all analyses
AR_MD_03	Missing DAA due to reasons other than lack of safety	0	Include in all analyses

Table 5-5 describes subject classification with regards to analysis sets:

Table 5-5 Subject classification

Analysis Set	PD ID that may cause subjects to be excluded	Non-PD (AR) ID that cause subjects to be excluded
RAN	M_INCL01_ICF not obtained	Not Randomized;
FAS	M_INCL01_ICF not obtained	Not in the RAN; Did not receive at least one study injection
SAF	M_INCL01_ICF not obtained	Did not receive at least one study injection
PPS	M_INCL01_ICF not obtained P_INCL02_Age less than 18 yrs, M_INCL03_Diabetes eligibility criteria, P_INCL04_No visual impairment (study eye), M_EXCL01_Confounding condition in study eye, M_EXCL02_Confounding concomitant medications or procedures in study eye, M_TRT01_Wrong IP administered, M_TRT02_Under-treatment during loading phase, M_TRT03_Over treatment, M_TRT04_Under-treatment after loading phase, M_OTH01_Masking process not followed, M_OTH02_Any other PD, P_WITH01_Treatment but consent withdrawn	Not in the FAS AR_EST_03, AR_ETD_03, AR_MD_01

5.6 Censoring rules for analysis

Protocol deviations (PDs) and analysis restrictions (ARs) that are considered to be critical for the subject evaluability regarding the primary and key secondary endpoints are described in Section 5.5.

The focus of the ARs is the identification of censoring related to the analysis of BCVA and q12w proportion as derived from DAA and described in Section 2.6.2. Censoring for DAA is only applied in case the underlying reason for a confounded DAA is assessed as '0'. Censoring of BCVA and DAA applies both to the year-1 analysis and year-2 analysis for the FAS, and only to the year-1 analysis for the PPS.

Table 5-6 summarizes the concepts of censoring for the key parameters BCVA and q12w-status/DAA applied to the two efficacy analysis sets, FAS and PPS, as well as the details for the timing of censoring for BCVA and DAA.

In case a subject has multiple PDs/ARs with impact on subject's evaluability the following rules are applied:

- A subject is excluded from an analysis set if at least one PD or AR with this consequence was identified (see Table 5-5). This rule is built on the concept of the medical assessment of the 'reason' which considers the reason of an earlier event to potentially also be the reason for following PDs or ARs.
- In case of multiple censoring time points censoring will be performed at the earliest.

Table 5-6 Censoring concepts for BCVA and DAA

Analysis Set	Censoring concept for BCVA	Censoring concept for DAA
FAS	Censoring of BCVA data after switch to alternative DME treatment in the study eye: imputation using the last observation collected prior to the start of alternative DME treatment (see section 2.5.3) No other censoring related to PDs or ARs.	M_TRT01_Wrong_IP_administered: censoring at the last valid DAA visit at or prior to the PD visit M_TRT02_Under-treatment during loading phase: censoring at baseline M_TRT03_Over treatment: censoring at the last valid DAA visit at or prior to the PD visit M_TRT04_Under-treatment after loading phase: censoring at the last valid DAA visit at or prior to the PD visit M_COMD01_Prohibited medication or procedure: censoring at the last valid DAA prior to the start of the prohibited medication or procedure AR_ETD_03: censoring at the last valid DAA visit at or prior to the PD visit AR_EST_03: censoring at the last valid DAA visit at or prior to the PD visit AR_MD_03: censoring at the last valid DAA prior to the missed visit Remark: The primary analysis of the q12w proportion as derived from DAA and described in section 2.6.2 applies censoring in case the underlying DAA is considered to be confounded by reasons other than lack of efficacy and/or safety. Based on the underlying time-to-'first-q8w-need' analysis, all information up to and including the censoring time-point contribute to the evaluation of the q12w status. **Censoring: subjects are considered to no longer be under risk for a q8w-need identification at later visits.** Censoring at baseline if above PD/AR occurred prior to Week 32. Censoring at Week 52 visit if subjects completed Week 52 without above PD/AR (only applies to Week 52 analysis)
PPS	Censoring of BCVA data after switch to alternative DME treatment in the study eye: imputation using the last observation	Similar to FAS

	collected prior to the start of alternative DME treatment (see section 2.5.3) M_COMD01_Prohibited medication or procedure: censor at the last observation collected prior to the start of the prohibited medication or procedure, imputation using the last observation collected prior to the start of prohibited medication or procedure	

6 References

Bretz F, Maurer W, Brannath W, et al (2009) A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine*; 28(4): 586-604.

Bretz, F., Maurer, W. and Maca J. (2014) Graphical Approaches to Multiple Testing. To appear as Chapter 14 in: *Clinical Trial Biostatistics and Biopharmaceutical Applications* (ed: Walter Young and Ding-Geng (Din) Chen), Taylor and Francis, Boca Raton

Clinical Development

RTH258/Brolucizumab

CRTH258B2301 / NCT03481634

A Two-Year, Three-Arm, Randomized, Double-Masked, Multicenter, Phase III Study Assessing the Efficacy and Safety of Brolucizumab versus Aflibercept in Adult Patients with Visual Impairment due to Diabetic Macular Edema (KESTREL)

Statistical Analysis Plan (SAP)

Author: [redacted]

Document type: SAP Documentation

Document status: Amendment V2.0 – clean version

Release date: 28-Oct-2021

Number of pages: 52

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
26-Jun-2020	Prior to DB lock	Creation of first version	N/A	NA
26-Nov-2020	Prior to DB lock	Creation of amendment 1	Reflect the change in the analysis timing to occur when all subjects complete Week 52 visits instead of the first 534 randomized subjects	Sections 2.1, 2.13
			Add analyses to cover potential impact of the COVID-19 pandemic	Sections 2.3.1, 2.5.4.2, 2.5.4.3, 2.6.4, 2.7.2.2, 2.7.2.3, 2.8.1
			Add back the summary tables of medical history and prior/concomittant medications, as per the protocol	Sections 2.3.3, 2.4.2
			Add clarifications for the definition of primary and supplementary estimands	Sections 2.5.4.1, 2.5.4.2
			Add clarifications for the q12w treatment status	Section 2.6.2
			Update testing strategy to add confirmatory superiority testing on efficacy endpoints	Section 2.7.2.1
			Add details on the determination of AESIs	Section 2.8.1.1
			Add section regarding additional imaging parameters for subjects with AESI in the study eye	Section 2.8.4.4
			Correct minor issues in the first version	
28-Oct-2021	Prior to Y-2 DB lock	Creation of amendment 2	Addition of COVID-19 impacted vs non-impacted analysis in Year-2 CSR	Section 2.5.4.3, Section 2.7.2.2, Section 2.8.1
			Clarification on analysis of Year-2 dosing regimen endpoints	Section 2.7.1, Section 2.7.2.4
			Updates from previous amendment as captured in the Year-1 CSR section 9.8.3 “changes in planned analysis“	Section 2.1.1, Section 2.10

Table of contents

List of abbreviations

ADA	Anti-drug antibody
AE	Adverse Event
ALP	Alkaline Phosphate
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AR	Analysis Restrictions
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Classification
BCVA	Best Corrected Visual Acuity
CI	Confidence Interval
cm	Centimeter
CF	Color Fundus
COVID-19	Coronavirus Disease 2019
CSFT	Central subfield thickness
CSFTns	Central Subfield Thickness-neurosensory retina
CSR	Clinical Study Report
CRS	Case Retrieval Strategy
CRC	Central Reading Center
DBL	Database Lock
DMC	Data Monitoring Committee
DME	Diabetic Macular Edema
DR	Diabetic Retinopathy
DRSS	Diabetic Retinopathy Severity Scale
eCRF	Electronic Case Report Form
eCRS	Electronic Case Retrieval Strategy
EDC	Electronic Data Capture
ETDRS	Early Treatment Diabetic Retinopathy Study
EoT	End of Treatment
EoS	End of Study
FA	Fluorescein Angiography
IA	Interim Analysis
IOP	Intraocular Pressure
IRT	Interactive Response Technology
IRF	Intraretinal Fluid
IVT	Intravitreal Treatment
kg	Kilogram
KM	Kaplan Meier
LOCF	Last observation carried forward
LSM	Least Square Means
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligrams

MICE	Multiple Imputation by Chained Wquations
mL	Millilitres
MMRM	Mixed Model for Repeated Measures
OCT	Optical Coherence Tomography
PD	Protocol Deviation
PDS	Programming Dataset Specifications
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD-OCT	Spectral Domain Optical Coherence Tomography
SE	Standard Error
SOC	System Organ Class
ULN	Upper Limit of Normal
q1	25th Percentile
q3	75th Percentile
q4w	Every 4 Weeks
q6w	Every 6 Weeks
q8w	Every 8 Weeks
q12w	Every 12 Weeks
SRF	Subretinal Fluid
TBL	Total Bilirubin
TEAE	Treatment-Emergent Adverse Event
TFLs	Tables, Figures, Listings
VA	Visual acuity
VFQ-25	Visual Functioning Questionnaire-25
WHO	World Health Organization

1 Introduction

The purpose of this Statistical Analysis Plan (SAP) is to describe the implementation of statistical analyses planned in the study protocol, and to provide detailed statistical methods that will be used for the Clinical Study Reports (CSRs) of study CRTH258B2301, at Week 52 (Year-1 CSR) and at Week 100 (Year-2 CSR).

Data will be analyzed according to the data analysis Section 9 of the study protocol which will be available in Appendix 16.1.1 of the CSR. Important information is given in the following sections and additional details will be provided, as applicable, in Appendix 16.1.9 of the CSR.

The SAP will be finalized before the interim database lock (DBL) for the primary analysis at Week 52. Any changes to the SAP after approval will be documented.

The following document was referenced while writing this SAP:

CRTH258B2301 Clinical Trial Protocol Final version 04 dated 08-Oct-2021

1.1 Study design

This is a randomized, double-masked, multi-center, active-controlled 3-arm study in patients with diabetic macular edema (DME) to evaluate the safety and efficacy of brolucizumab 3 mg and 6 mg against the active control aflibercept 2 mg.

Approximately 700 patients will be screened in order to randomize a total of approximately 534 patients (178 per arm) in a 1:1:1 ratio in one of the three treatment arms:

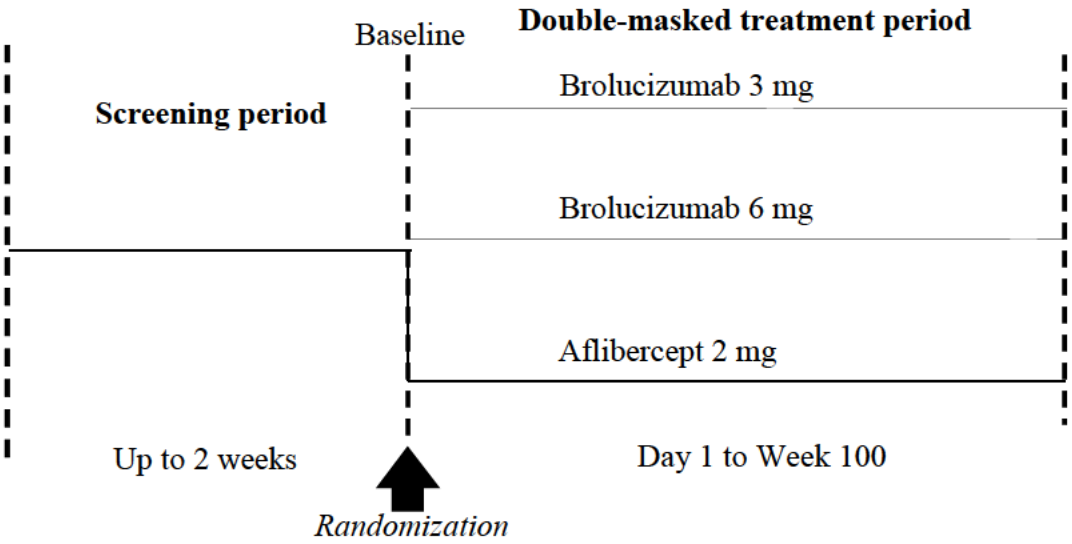
- Brolucizumab 3 mg: 5 × q6w loading then q12w/q8w maintenance
- Brolucizumab 6 mg: 5 × q6w loading then q12w/q8w maintenance
- Aflibercept 2 mg: 5 × q4w loading then q8w maintenance

At the baseline visit, all eligible patients will be randomized via Interactive Response Technology (IRT) to one of the treatment arms. Stratification for Japanese ethnicity (Japanese vs. non-Japanese) will be considered.

Since the treatment schedule is different for brolucizumab and aflibercept treatment arms the following will be applied to ensure masking:

- In addition to visits every 4 weeks for all patients for 2 years, extra visits are scheduled at Weeks 6 and 18 for all treatment arms
- The patients will receive active/sham injection at each protocol visit except at Weeks 20, 28 and 100 visits (no scheduled treatment for any arm)
- Disease activity assessment will be performed for all arms at pre-specified visits
- To fulfil the double-masking requirement, the investigational site will have masked and unmasked staff

Figure 1-1 Study design



There will be two treatment phases for IVT injections with different timing for brolucizumab and aflibercept treatment arms:

Loading Phase:

Brolucizumab 3 mg or 6 mg: In the loading phase, treatment with brolucizumab 3 mg or 6 mg will occur every 6 weeks for five (5) consecutive injections (Baseline, Weeks 6, 12, 18 and 24). To preserve the masking, the patients assigned to this regimen will receive sham injection on Weeks 4, 8 and 16.

Aflibercept 2 mg: In the loading phase, treatment with aflibercept 2 mg will occur every 4 weeks for five (5) consecutive injections (Baseline, Weeks 4, 8, 12 and 16). To preserve the masking, the patients assigned to aflibercept arm will receive sham injection on Weeks 6 and 18.

Maintenance Phase:

The treatment interval during the maintenance phase will be as follows:

Brolucizumab 3 mg or 6 mg: From Week 24 onwards, patients will be scheduled to receive one injection of brolucizumab 3 mg or 6 mg every 12 weeks. If, however, disease activity is identified by the evaluating/masked investigator at pre-specified visits, the patient will be assigned to receive treatment every 8 weeks (please refer to protocol for 'Evaluation of Disease Activity'). A disease activity assessment will also be performed at Week 96 to document the adequacy of the q12w treatment schedule at the end of the 2 year follow-up but will not be entered into IRT and will have no effect on the patient's treatment schedule.

Aflibercept 2 mg: From Week 16 onwards, patients will receive one injection of aflibercept 2 mg every 8 weeks (first injection after Week 16 to be given at Week 24) until Week 96 visit. Disease activity assessments will be conducted at pre-specified visits by the evaluating/masked investigator for masking purposes and will not influence the treatment interval.

The primary analysis will be based on the Week 52 data, i.e. all data up to and including Week 52 (see Section 2.1).

Patients will remain in the study and will continue to receive masked treatment through the planned duration (100 weeks) to allow for further masked evaluation of efficacy and safety. Treatment masking of individual patients will remain intact for all patients, investigators and selected staff from the Sponsor who have contact with patients or investigators or those who are involved in the direct conduct of the study until the final database lock has occurred.

1.2 Study objectives and endpoints

Study objectives and related endpoints are described in Table 1-1 below.

Table 1-1 Objectives and related endpoints

Objective	Endpoint
Primary	
To demonstrate that brolucizumab is noninferior to aflibercept with respect to the visual outcome after the first year of treatment	Change from baseline in best-corrected visual acuity (BCVA) at Week 52
Secondary	
To demonstrate that brolucizumab is noninferior to aflibercept with respect to visual outcome during the last 3 months of the first year of treatment	Change from baseline in BCVA averaged over a period Week 40 to Week 52
To estimate the proportion of patients treated at q12w frequency with brolucizumab	Proportion of patients maintained at q12w up to Weeks 52 & 100
To estimate the predictive value of the first q12w cycle for maintenance of q12w treatment with brolucizumab	Proportion of patients maintained at q12w up to Weeks 52 & 100, within those patients that qualified for q12w at Week 36
To evaluate the functional and anatomical outcome with brolucizumab relative to aflibercept	Change from baseline by visit up to Week 100 in BCVA and in parameters derived from SD-OCT, Color fundus photography and Fluorescein angiography
To evaluate the effect of brolucizumab relative to aflibercept on the Diabetic Retinopathy status	Change in ETDRS Diabetic Retinopathy Severity Scale (DRSS) score up to Week 100
To assess the safety and tolerability of brolucizumab relative to aflibercept	Incidence of Ocular and Non-ocular AEs, vital signs and laboratory values up to Week 100
To evaluate the effect of brolucizumab relative to aflibercept on patient-reported outcomes (VFQ-25)	Change in patient reported outcomes (VFQ-25) total and subscale scores from baseline up to Week 100

2 Statistical methods

2.1 Data analysis general information

The primary safety and efficacy analysis will be based on the Week 52 data, i.e. all data up to and including Week 52. This analysis will be performed when all randomized subjects have completed their Week 52 visit or terminated the study before Week 52, while subjects continue to receive masked treatment through the planned study duration of 100 weeks.

A second planned interim analysis may be performed by locking the Week 76 data in case of regulatory request of supplemental data to be submitted during the review period.

The statistical analysis will be performed by Novartis using SAS Version 9.4 or above.

Continuous variables will be summarized using the number of observations, mean, standard deviation, standard errors (SE), median, quartiles, minimum and maximum values. Categorical variables will be summarized with number of observations, the number of observations for each category and the corresponding percentage. Where appropriate, 2-sided 95% confidence intervals (CIs) for point estimates of the mean or proportion will be provided. For the treatment difference brolucizumab – aflibercept, point estimates, 95% CIs will be provided as appropriate unless otherwise specified.

All the analyses listed in the SAP that correspond to data collected during the 2nd year of the study (post-Week 52) will be part of the end of study (year-2) CSR only. For the year-2 CSR, efficacy endpoints specific to the first year, i.e., up to the Week 52 visit (e.g., the change from baseline in BCVA averaged over Week 40 to Week 52) will not be analyzed and reported again; analyses by visit and assessments based on cumulative data (e.g., incidence of AEs) will include data from baseline up to the end of study.

2.1.1 General definitions

Study drug and study treatment

Study drug refers to brolucizumab 3 mg, brolucizumab 6 mg and aflibercept 2 mg IVT injections.

Study treatment refers to study drug or sham injections.

Study day

Day 1 is defined as the date of first administration of study treatment. Study day is defined as the number of days relative to the date of first dose of study treatment (Day 1).

Therefore, for a particular date, study day will be calculated as follows:

- for dates on or after the date of first administration of study treatment:

  Study day = Assessment date – Date of first administration of study treatment + 1;
- for dates prior to the date of first administration of study treatment:

  Study day = Assessment date – Date of first administration of study treatment.

Baseline

The baseline value is defined as the last assessment performed prior to administration of the first dose of study treatment.

All data collected after first study treatment are defined as post-baseline.

End of study day mapping

The end of study (EoS) date is the date when a subject completes or discontinues the study.

For reporting data by visit in outputs, the end of study visit will be allocated to the actual (reported) visit number. If end of study date is not on a scheduled visit, then the EoS visit will be allocated, based on study day, to the closest future scheduled study visit.

End of treatment day mapping

The "Date of Last Exposure" is the date of the last study treatment on or prior to the end of treatment (EoT) date.

For reporting data by visit in outputs, the EoT date will be allocated to the actual (reported) visit number. If EoT date is not on a scheduled visit, then the EoT date will be allocated, based on study day, to the closest future scheduled study visit.

Unscheduled visits

Data collected at unscheduled visits will not be used in 'by-visit' tabulations or graphs , but will be included in analyses based on all post-baseline values such as last observation carried forward (LOCF) imputation, and summary of maximum letter loss in BCVA from baseline. These data would not be used in case of analyses with mixed model for repeated measures (MMRM). Only values at scheduled visits (observed or imputed by LOCF) will be considered to average BCVA change from baseline over a given period. Moreover, given unscheduled visits will not be active treatment visits, IOP measurements at unscheduled visits will not be considered as pre-injection IOP measurements, hence will not be used to identify subjects with pre-injection IOP >30mmHg.**Missing and implausible dates**

The general approach to handling missing dates is shown in Section 5.1.

2.2 Analysis sets

The **All Enrolled Set** includes all subjects who signed informed consent. This analysis set will be used to summarize subject disposition.

The **Randomized Set** will consist of all randomized subjects. Subjects are considered randomized when they have been deemed eligible for randomization by the investigator and given a randomization number. Subjects will be analyzed according to the treatment assigned to at randomization.

The **Full Analysis Set** (FAS) includes all randomized subjects who receive at least one IVT injection of the study treatment. The full analysis set will serve as the primary analysis set for all efficacy analyses. Subjects will be analyzed according to the treatment assigned to at randomization.

Supportive analyses of the primary and key secondary endpoints will include analyses using the **Per Protocol Set** (PPS). PPS is a subset of the FAS and will exclude or censor subjects with

important protocol deviations (PDs) and analysis restrictions (ARs) that are expected to majorly affect the validity of the assessment of efficacy and/or safety at Week 52, including for e.g. lack of compliance (including missed treatments and treatment misallocation), missing data, prohibited concomitant medication and deviations from inclusion/exclusion criteria. Confounded data or discontinuation from study treatment due to lack of efficacy and/or safety do not constitute a reason for exclusion from the PPS.

Before the Week 52 database lock the relevant protocol deviations will be identified at the subject level in the database. After the Week 52 database lock, analysis restrictions will be derived in the analysis database. Censoring applied in relation to the specific PDs / ARs will be specified as well.

The FAS will be the analysis set for the primary estimand as defined in Table 2-1. However, when assessing the robustness of the overall efficacy conclusions, considerations will be given to the analysis based on the primary estimand using FAS and the supplementary estimand (see Table 2-1) using PPS, i.e., similar conclusions on non-inferiority based on both estimands are expected. Inconsistencies in key efficacy study results between the FAS and PPS will be examined and discussed in the clinical study report (CSR). PPS will not be used in any analysis related to year-2 CSR.The **Safety Analysis Set** (SAF) will include all subjects who receive at least one IVT injection. Subjects in the safety analysis set will be analyzed according to the treatment arm from which they received the majority of treatments up to and including Week 48.

Prior to locking the database for the primary analysis at Week 52 and breaking the masked treatment assignment code, the relevant important protocol deviations will be identified as specified in Section 5.6. The corresponding identifications at the subject level including data exclusion from PPS and censoring will be captured in the database. Analysis Restrictions (non protocol deviations) will be identified by programming (as specified in the programming specification document) independently to the treatment arm.

Rules of exclusion criteria of analysis sets are in Appendix Section 5.5.

For the primary analysis performed when the first 534 randomized subjects have completed their Week 52 visit or terminated the study before Week 52, footnotes will clarify that the analysis sets considered for the outputs are not considering all randomized subjects.

2.2.1 Subgroups of interest

The subgroups of interest are specified below:

- Age category (<65, ≥65 years)
- Gender (male, female)
- Diabetes type (Type 1, Type 2)
- Baseline HbA1c (<7.5, ≥7.5%)
- Baseline BCVA categories (≤65, >65 letters)
- Duration of DME since the primary diagnosis (≤3, >3-<12, ≥12 months)
- DME type (focal, diffuse) as per central reading center (CRC)
- Baseline central subfield thickness (CSFT) (<450, ≥450-<650, ≥650 µm)
- Baseline status of intraretinal fluid (IRF) (presence, absence)

- Baseline status of subretinal fluid (SRF) (presence, absence)
- Ethnicity (Japan, non-Japan)

Subgroup analyses will be performed in the year-1 CSR for the primary and key secondary efficacy variables only (as defined in Section 2.5.1 and Section 2.6.1), using the primary analysis approach. More details can be found in Section 2.5.4 and Section 2.6.2.

2.3 Subject disposition, demographics and other baseline characteristics

2.3.1 Subject disposition

The following summaries will be included in the disposition table considering all enrolled subjects: number and percent of subjects who were enrolled into the study, treated, completed the study (Week 52/Week 100), discontinued the study (prior to or at Week 52/Week 100) (including reasons for discontinuation) and discontinued from study treatment (prior to or at Week 52/Week 96) (including reasons for discontinuation).

The number and percentage of subjects who discontinued the study and who discontinued study treatment will be presented by treatment arm and study visit. The number and percentage of subjects treated by site and treatment arm will be presented.

A listing of subjects who discontinued from the study and/or treatment early will be provided by treatment arm. The listing will identify the visits completed and when the study or treatment was discontinued including the corresponding reasons.

Subjects who signed an informed consent form and who were subsequently found to be ineligible prior to randomization will be considered a screen failure. Screen failure information will not be summarized but only listed.

Number and percentage of subjects who were excluded (i.e. not evaluable) from each of the SAF, FAS, and PPS will be presented using the randomized analysis set. A listing of subjects along with the analysis set that they were excluded from and the corresponding reasons will also be presented.

Number and percentage of subjects with important protocol deviations (PD) and analysis restrictions (AR) will be presented by treatment arm and deviation/restriction category. Due to the COVID-19 pandemic, higher number of PDs are expected. To evaluate the PDs that occurred due to COVID-19, the number and percentage of subjects with PDs that occurred due to COVID-19 outbreak will also be provided by deviation category and treatment arm. A listing of all ARs and PDs will be provided by treatment arm and subject, including the information if the AR/PD leads to the subject exclusion from an analysis set and the relationship to COVID-19.

2.3.2 Demographics and baseline characteristics

Demographics and baseline characteristics will be summarized with descriptive statistics for the FAS by treatment arm and overall. Demographic characteristics will include age, gender, race, and ethnicity. The summary of baseline ocular characteristics will be presented for the study eye only and listed separately for the study eye and the fellow eye.

Ocular baseline characteristics include:

- Study eye selection (left eye OS or right eye OD),
- Diabetes type (Type 1, Type 2),
- Duration of DME since the primary diagnosis as a continuous variable and using categories (<3, >3-<12, ≥12 months),
- Macular edema type (focal, diffuse) as per CRC,
- Baseline BCVA as continuous variable and using categories (≤65, >65 letters, and <60, ≥60-≤70, >70 letters),
- Baseline HbA1c as continuous variable and using categories (<7.5, ≥7.5%),
- Baseline CSFT (<450, ≥450-<650, ≥650 µm),
- Baseline status of IRF (presence, absence),
- Baseline status of SRF (presence, absence),
- Baseline ETDRS DRSS score using categories (12-point scale, see Table 2-4).

Duration of DME since diagnosis (months) will be derived as [(first dose date – diagnosis start date + 1)/(365.25/12)]. In case of partial dates, the imputation rule is specified in Section 5.1.4.

Other relevant baseline information will be listed and summarized with descriptive statistics as appropriate.

No tests for differences in demographics and baseline characteristics between treatment arms will be performed. Potential differences will be assessed based on clinical relevance.

2.3.3 Medical history

Medical history and current medical conditions will be summarized and listed for ocular (study eye) and non-ocular events.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment exposure

Extent of exposure to study treatment is calculated as the number of study treatment injections received.

Descriptive statistics for exposure to study treatment will be provided for the Safety set.

The following summaries will be presented:

- Overall number of treatments cumulatively for the period baseline to Week 48/52 (Week 96) including separate analysis for the loading phases, i.e. up to Week 16 (last treatment of the 5 x q4w loading of aflibercept) and up to Week 24 (last treatment of 5 x q6w loading of brolucizumab), and maintenance phase, using the following categories: active and sham IVT injections, active only, sham only
- Treatment exposure by visit: the number and percentage of subjects who received active IVT injections of study treatment, sham injections, missed a treatment (active and sham) and missed visits will be presented by treatment arm and visit

- Frequency of all observed dosing patterns from baseline to Week 52 (Week 100), differentiating between active and sham treatments, missed study treatments and wrong study treatments
- Brolucizumab treatment allocation by visit from Week 32 onwards: number and percentage of subjects on q12w and q8w at each visit, including number of subjects switched from q12w to q8w.

Exposure data will be listed for all treatment arms.

2.4.2 Prior and concomitant medications

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Concomitant medications are defined as medications received after the start of study treatment including those already started prior to the start of the study treatment.

Prior and concomitant medications will be coded according to the WHO Drug Reference List dictionary, with Anatomical Therapeutic Classification (ATC) class and preferred term.

Ocular and non-ocular prior and concomitant medications will be summarized and listed by ATC class and preferred term (PT) by treatment arm. Ocular medications will be listed for the study eye and the fellow eye separately.

Anti-VEGF medications will be summarized by ATC class and preferred term for systemic route, the study eye and the fellow eye separately by treatment arm.

Ocular concomitant non-drug therapies and procedures will be summarized for the study eye only. Both ocular and non-ocular concomitant non-drug therapies and procedures will be listed.

In the summary tables, data collected after the subject discontinued study treatment and started alternative DME treatment in the study eye will be censored (from the day the subject started alternative DME treatment onwards).

2.5 Analysis of the primary and first key secondary endpoints

2.5.1 Primary and first key secondary endpoints

The primary endpoint is the change from baseline in BCVA at Week 52 in the study eye (ETDRS letters).

The first key secondary endpoint is the average change from baseline in BCVA over the period Week 40 through Week 52 in the study eye. For each subject, this endpoint is defined as the average of the changes from baseline to Weeks 40, 44, 48 and 52.

The motivation for the choice of this endpoint is that, averaging the BCVA values over Week 40 to Week 52 will address both random fluctuations and potential trough and peak values during the different treatment cycles (q8w and q12w).

The primary analysis of the primary and first key secondary endpoints will be based on the FAS with last observation carried forward (LOCF) imputation of missing or censored BCVA values.

The primary estimand for the primary endpoint includes the following components:

- Population: Subjects with visual impairment due to DME as per the inclusion/exclusion criteria
- Endpoint: The primary endpoint is the change from baseline in BCVA at Week 52. BCVA will be assessed by the masked investigator using ETDRS-like charts at an initial distance of 4 meters.
- Treatment of interest: The randomized study treatment (brolucizumab or aflibercept)
- The handling of the remaining intercurrent events as follows:
  - Study discontinuation due to any reason: data imputed with LOCF
  - Treatment discontinuation due to any reason: use all the data
  - Data after the start of alternative DME treatment will be censored
  - Summary measure: Difference in the change from baseline in BCVA at Week 52 between brolucizumab and aflibercept treatment arms.

The primary estimand for the first key secondary endpoint has similar components.

2.5.2 Statistical hypothesis, model, and method of analysis

The objective related to the primary and first key secondary endpoints is to demonstrate non-inferiority of brolucizumab to aflibercept with respect to change from baseline in BCVA, considering a margin of 4 ETDRS letters.

Let:

B3 = Brolucizumab 3 mg - 5 x q6w loading then q12w/q8w maintenance

B6 = Brolucizumab 6 mg - 5 x q6w loading then q12w/q8w maintenance

A = Aflibercept 2 mg - 5 x q4w loading then q8w maintenance

The following non-inferiority hypotheses are related to a non-inferiority margin of 4 letters:

$\mathbf{H_{01}}: \mu_{B6} - \mu_A \leq -4$ letters vs. $\mathbf{H_{A1}}: \mu_{B6} - \mu_A > -4$ letters

$\mathbf{H_{02}}: \varphi_{B6} - \varphi_A \leq -4$ letters vs. $\mathbf{H_{A2}}: \varphi_{B6} - \varphi_A > -4$ letters

$\mathbf{H_{03}}: \mu_{B3} - \mu_A \leq -4$ letters vs. $\mathbf{H_{A3}}: \mu_{B3} - \mu_A > -4$ letters

$\mathbf{H_{04}}: \varphi_{B3} - \varphi_A \leq -4$ letters vs. $\mathbf{H_{A4}}: \varphi_{B3} - \varphi_A > -4$ letters

where μ_{B6} μ_{B3} and μ_A are the corresponding unknown true mean changes from baseline in BCVA at Week 52; φ_{B6}, φ_{B3} and φ_A are the corresponding unknown true mean changes from baseline in BCVA averaged over the period Week 40 to Week 52.

Based on the FAS, the above hypotheses will be tested via an analysis of variance (ANOVA) model. The model will include treatment, baseline BCVA (≤65, >65 letters) and age category (<65, ≥65 years) as factors. Two-sided 95% confidence interval (CI) for the least square mean (LSM) difference (brolucizumab - aflibercept) will be presented in letters. Non-inferiority will be considered established if the lower limit of the corresponding 95% CI is greater than -4 letters.

P-value for treatment comparison (2-sided) and p-value for non-inferiority (4 letter margin) (1-sided) will be presented.

These 4 hypotheses will be tested sequentially in the order of their numbering (H_{An}, n=1, 2, 3, 4), i.e., confirmatory testing of the second, third or fourth hypotheses requires rejection of each preceding null hypothesis. In this setting, each hypothesis will be assessed at a one-sided significance level of 0.025.

The primary estimand and other supplementary estimands of interest are described in Table 2-1 below, together with their key attributes. The same approach for non-inferiority assessment in change from baseline in BCVA at Week 52 and average change from baseline in BCVA over the period Week 40 through Week 52 will be applied to any supplementary estimand.

Table 2-1 Primary and supplementary estimands

Estimand	Estimand definition	Analysis set	Statistical methods (Including strategy for imputation/replacement of missing/censored data)
Primary estimand	Difference in change from baseline in BCVA at Week 52 excluding the effect of switching to alternative DME medication in the study eye	FAS	Analysis of variance (ANOVA) model including terms for treatment, baseline BCVA (≤65, >65 letters) and age category (<65, ≥65 years), and using LOCF imputation/replacement for missing/censored data.
Supplementary estimand	Difference in change from baseline in BCVA at Week 52 for subjects adhering to the protocol as per the PPS definition	PPS	ANOVA model as per the primary estimand. LOCF imputation/replacement for missing/censored data

For additional information on data handling related to intercurrent events, see Section 5.5 and Section 5.6.

2.5.3 Handling of missing values/censoring/discontinuations

Missing BCVA values will be imputed by LOCF as a primary approach. Observed values from both scheduled and unscheduled visits will be used for the LOCF imputation. For subjects with no post-baseline BCVA value, the baseline value will be carried forward.

For subjects who discontinue study treatment but continue in the study, data collected after the switch to alternative DME treatment in the study eye (see Table 2-2) will be censored for the primary analysis. Censored data will be replaced using LOCF with the last observation collected prior to the start of alternative DME treatment in the study eye.

Table 2-2 Potential alternative DME treatment in the study eye

- Ranibizumab
- Aflibercept
- Bevacizumab (off-label use)
- Laser photocoagulation, previous standard of care still being used as mono- or combination therapy with anti-VEGF
- Intraocularly administered steroids:
    - Dexamethasone

- Fluocinolone acetonide

From an estimand perspective, the main focus is to adequately reflect in the analysis unfavorable study outcomes related to the treatment (e.g. lack of efficacy, safety problems).

The LOCF approach is expected to be sensitive to an early study termination due to lack of efficacy assuming that such lack of efficacy is reflected in the last observed BCVA measurement. In case of the use of alternative treatment for the underlying disease (DME), data collected after the start of such a treatment would be censored. LOCF will then be based on the last value prior to the start of this treatment, again expecting that this value would reflect the negative BCVA outcome under study treatment. In case of missing data due to lack of safety/tolerability with impairment of the function of the study eye the LOCF method would also provide a sensitive approach to capture such an unfavorable outcome.

In case of missing data occurring independently of the response to study treatment, the LOCF approach assumes stability which seems to be adequate based on historical data both for the maintenance treatment phase (i.e. stabilization of BCVA) and also in case of the absence of any treatment effect with an average natural disease progression in terms of BCVA of only 1-2-letter loss over 1 year. In case of an early study termination during the loading phase occurring independently of the response to study treatment, the LOCF method will result in a conservative estimate potentially underestimating the true outcome.

LOCF is an established method within the assessment of efficacy of anti-VEGF treatments in terms of BCVA outcome. Non-inferiority studies should follow the main design features (primary variables, the dose of the active comparator, eligibility criteria, etc.) as the previously conducted superiority trials in which the active comparator demonstrated clinically relevant efficacy. The primary endpoint in aflibercept Phase III studies VIVID and VISTA was the BCVA change from Baseline to Week 52 with missing data imputed based on LOCF. Based on those studies, the percentage of missing data regarding BCVA is not considered critical (<10%) which limits the impact of the missing data imputation method.

2.5.4 Sensitivity and supportive analyses

2.5.4.1 Sensitivity analyses on the primary estimand

Sensitivity to the statistical model and imputation assumptions from the primary estimand will be considered, using the primary analysis set (FAS) only.

An alternative method of handling missing/censored values as described below may be considered to assess the robustness of the hypothesis testing resulting from the primary analysis described in Section 2.5.2:

- Mixed model for repeated measures (MMRM) assuming missing at random (MAR) using observed data only (including censoring of BCVA values collected after the start of alternative DME treatment). The MMRM will include treatment, visit, baseline BCVA category, age category and treatment by visit interaction as fixed-effect terms, and visit as a repeated measure. An unstructured covariance matrix will be used to model the within-subject error. For the MMRM analysis:
  - The treatment difference brolucizumab - aflibercept at Week 52 will be estimated using the LSM and 95% CI.

- For the endpoint of average change from baseline over the period Week 40 through Week 52, a SAS code using the ESTIMATE statement in PROC MIXED will be provided in the programming specification document to obtain the LSM estimate and 95% CI for the corresponding treatment difference.
- If an MMRM model with an unstructured covariance matrix does not converge, a more restricted covariance matrix can be considered in the following order until convergence is reached: compound symmetry (CS), first-order autoregressive (AR), Toeplitz (TOEP), and variance components (VC).

In this analysis, data collected after the switch to alternative DME treatment in the study eye will be censored.

Other sensitivity analyses on the primary estimand might be considered, such as tipping point analysis or multiple imputation by chained equations (MICE) method.

2.5.4.2 Supportive analysis using a supplementary estimand

Supplementary estimand on the PPS:

The target population, the primary endpoint, the treatment of interest and the summary measure of the supplementary estimand are the same as for the primary estimand. The handling of the intercurrent events for the supplementary estimand can be found in Table 5-6 for the PPS population.

The supportive analysis on the supplementary estimand will apply the same LOCF/ANOVA method as for the primary estimand.

Supplementary estimand to assess the impact of COVID-19:

Another supplementary estimand might be defined to assess the impact of intercurrent events associated with study treatment discontinuation due to COVID-19 on the study conclusions. For subjects who discontinue study treatment due to COVID-19 but continue in the study, data collected after the treatment discontinuation will be censored for the analysis. Censored data will be replaced using LOCF with the last observation collected prior to the study treatment discontinuation. This analysis will be conducted on the FAS if at least 5% of subjects discontinued treatment due to COVID-19.

2.5.4.3 Summary statistics and subgroup analysis

Summary statistics:

- Descriptive statistics of BCVA primary and first key secondary endpoints will use observed data and primary analysis set (FAS), with and without censoring data after use of alternative DME treatment in the study eye.

Subgroup analyses will be conducted to assess the consistency of treatment effect across various subgroups of interest as described in Section 2.2.1. They will be conducted using the framework for the primary estimand only (FAS with censoring of data collected after use of alternative DME treatment in the study eye, and missing/censored values imputed/replaced using LOCF):

- Subgroup analyses will be conducted using the same model and analysis strategies described for the primary and first key secondary endpoints but fitted by category of each of the subgroups. Subgroup variables that are used as fixed effects in the model will be removed from the model statement for the subgroup analysis.
- In case of analyses on subgroups with extremely imbalanced sample sizes, the subgroup levels can either be combined, if appropriate, or the extremely small subgroup will be excluded while fitting the analysis model.
- The point estimate and 95% CI for the between treatment difference for each subgroup will be presented using forest plots.

Subgroup analyses to evaluate impact of COVID-19 pandemic:
As per internal guidance, a sensitivity analysis related to the exposure of subjects to COVID-19 will be conducted. The definition of start and end dates by geographical areas to be used for the sensitivity analysis are as indicated below:

Region/Country	Start Date	End Date
China	01-Jan-2020	End date has not yet been defined
South Korea	20-Feb-2020	End date has not yet been defined
Japan	21-Feb-2020	End date has not yet been defined
Italy	23-Feb-2020	End date has not yet been defined
Rest of the World	01-Mar-2020	End dates have not yet been defined

In the year-1 CSR, non-exposed subjects to COVID-19 are defined as subjects who:

- completed Week 52 visit prior to the pandemic start date,
- or withdrew the study prior to the pandemic start date,
- or withdrew treatment and started alternative DME treatment prior to the pandemic start date.

Exposed subjects to COVID-19 are therefore defined as subjects who:

- did not complete Week 52 visit prior to the pandemic start date (while remaining in the study at the time of pandemic start date),
- or withdrew the study on or after the pandemic start date,
- or withdrew treatment and started alternative DME treatment on or after the pandemic start date,
- or withdrew treatment before the pandemic start date and did not start alternative DME treatment before the pandemic start date, while remaining in the study.

Furthermore, impacted subjects to COVID-19 pandemic were defined as subjects who:

- were exposed to COVID-19 as per the above definition,
- and missed at least one active injection due to COVID-19.

Non-impacted subjects to COVID-19 are therefore defined as subjects who:

- were not exposed to COVID-19 as per the above definition,
- or were exposed to COVID-19 but did not miss any active injection due to COVID-19.

In the year-1 CSR, subgroup analyses will be conducted using the same model and analysis strategies described for the primary and first key secondary endpoints in the exposed and non-exposed subgroups, and in the impacted and non-impacted subgroups. In addition,

demographics and baseline characteristics will be summarized for exposed and non-exposed subjects, and for impacted and non-impacted subjects.

Furthermore, in the year-2 CSR only, the impacted subjects to COVID-19 pandemic are defined as subjects who:

- missed at least one active injection due to COVID-19,
- or discontinued study/study treatment due to COVID-19,
- or had reported COVID-19 infection (including suspected as per PTs in the PDS).

Non-impacted subjects to COVID-19 are therefore defined as subjects who:

- did not miss any active injection due to COVID-19,
- and didn't discontinue study/study treatment due to COVID-19,
- and didn't have reported COVID-19 infection (including suspected as per PTs in the PDS).

In the year-2 CSR, subgroup analyses will be conducted using the same model and analysis strategies described for the secondary endpoints in the impacted and non-impacted subgroups (see Section 2.7.2.2). Demographics and baseline characteristics will be summarized for impacted and non-impacted subjects.

2.6 Analysis of additional key secondary endpoints

2.6.1 Additional key secondary endpoints

Additional key secondary endpoints are:

- Proportion of subjects maintained at q12w up to Week 52 (for brolucizumab treatment arms only)
- Proportion of subjects maintained at q12w up to Week 52, within those subjects that qualified for q12w at Week 36 (for brolucizumab treatment arms only)

2.6.2 Statistical hypothesis, model, and method of analysis

No hypothesis will be tested for the additional key secondary efficacy endpoints.

Following the estimand concept, in which consequences of lack of efficacy and/or lack of safety need adequate reflection in the efficacy estimates, the primary approach to derive the proportion of subjects with a positive q12w treatment status will be the "**efficacy/safety approach**", conducted using the FAS as described below.

The estimate for the proportion of subjects with a positive q12w treatment status will be derived from Kaplan-Meier (KM) time-to-event analyses for the event 'first q8w-need', imputing the 'q8w-need' in case of missing or confounded data attributable to lack of efficacy and/or lack of safety.

The proportion of subjects with a positive q12w treatment status will be derived as follows requiring 'sufficient duration of effect' (as assessed by q8w-need) together with 'sufficient efficacy and safety':

- For the ′sufficient duration of effect′ requirement subjects will need to have the status of ′q8w need =no′ at Weeks 32, 36 and 48 unless the ′q8w need = yes′ is confounded by reasons other than lack of efficacy and/or safety (see censoring details below)
- The requirement regarding ‘sufficient efficacy and safety’ will be addressed by considering subjects – even without an explicit ‘q8w-need = yes’ – as having a negative q12w status in case any of the following confounding factors is attributable to lack of efficacy and/or lack of safety of the study treatment: early treatment/study discontinuation, missed DAA. The q8w-need assessment is imputed as “Yes” at the DAA visit following early treatment/study discontinuation due lack of efficacy and/or lack of safety of the study treatment (applies to both missing and non-missing DAAs).

Intercurrent events associated with missing or confounded data regarding the q12w treatment status that are attributable to reasons other than lack of efficacy and/or safety are described below, together with the corresponding data handling strategies:

- Early treatment/study discontinuation: censoring at the last valid DAA visit while on treatment/study
- Single missed visit with a relevant DAA: censoring at the last valid DAA prior to the missed visit
- Prohibited concomitant medications/procedures: censoring at the last valid DAA prior to the corresponding application
- Discrepancy between DAA by investigator and the actual treatment received: censoring at the corresponding visit
- Other treatment allocations/applications deviating from the concept of ‘treatment allocation according to disease activity’: censoring at the last valid DAA at or prior to the deviating visit.

Censoring rules related to the q12w treatment status analysis are described in Section 5.6.

The proportion of subjects with a positive q12w treatment status at Week 52 will be presented together with two-sided 95% confidence intervals (see Section 5.4.2.3).

The outcome of the Kaplan-Meier analysis will be presented graphically by the estimated q12w-probability over time, i.e. at each DAA visit.

While for the analysis of the overall q12w proportion all subjects in the FAS will be considered, the analysis of the proportion of subjects maintained at q12w up to Week 52, within those subjects that qualified for q12w at Week 36, is based on the subset of FAS subjects with no identified q8w-need at Week 32 and Week 36. For this subset of subjects a valid Week 36 DAA is required, while missing the Week 32 assessment is considered as no q8w-need.

2.6.3 Handling of missing values/censoring/discontinuations

The details regarding handling of missing values and discontinuations, including the timing of censoring within the time-to-event analyses for the event ‘first q8w-need’, are specified in the previous section.

Remark: Subjects without any valid DAA are considered censored at baseline for the overall q12w proportion and for the analysis of the predictive value of the first q12w cycle.

From an estimand perspective, the impact of failing study completion according to the protocol due to lack of efficacy/safety is considered adequately reflected by a negative q12w-status.

2.6.4 Supportive analyses

Considerations around the occurrence of potential confounding effects impacting the assessment of the q12w treatment status at Week 52 lead to the development of different estimating approaches. Supportive analyses will be performed on the FAS using alternative methods of handling missing or confounded data:

- **'Efficacy only' approach**: approach with 'q8w-need' imputation only in case the reason for a missing or confounded q12w status is attributable to lack of efficacy of the study treatment. In case of a corresponding safety reason the subject is censored at the last valid DAA.
- **'As observed' approach:** analysis without 'q8w-need' imputation.

Additionally, analyses described in Section 2.6.2 conducted using the FAS will be repeated using the PPS to assess the consistency of the assessment of the q12w proportions when looking only at those subjects who adhere to the protocol.

Subgroup analyses will be conducted as well to assess the consistency of the assessment of the q12w proportions across various subgroups described in Section 2.2.1, considering the FAS only and the efficacy/safety approach only. It will include assessment of the impact of COVID-19 pandemic with the subgroup analyses of exposed/non-exposed subjects and impacted/non-impacted subjects as defined in Section 2.5.4.3.

In addition, potential confounding effects related to COVID-19 impacting the assessment of the q12w treatment status at Week 52 led to the development of a supplementary estimand. The analysis will be repeated using the FAS and the efficacy/safety approach only, after excluding subjects whose DAA assessment has been impacted by COVID-19 (e.g., DAA not performed due to COVID-19 site impact).

2.7 Analysis of secondary efficacy endpoints

2.7.1 Secondary efficacy endpoints

Secondary efficacy endpoints related to BCVA, dosing regimen, anatomy or status of diabetic retinopathy are listed below. All endpoints that consider data from post-Week 52 visits will not be part of the primary analysis at Week 52. These endpoints will only be analyzed in the year-2 CSR (Week 100).

[redacted]

Secondary efficacy endpoints based on BCVA:

- Change from baseline in BCVA at each visit up to Week 100.

- Average change from baseline in BCVA over the period Week 88 to Week 100 (year-2 analysis only). For each subject this endpoint is derived as the average of the changes from baseline to Weeks 88, 92, 96, 100.
- Average change from baseline in BCVA over the period Week 4 to Week 52/100. For each subject this endpoint is derived as the average of the changes from baseline to each post-baseline visit between Week 4 and Week 52/100.
- Average change from baseline in BCVA over the period Week 20 to Week 52/100 and Week 28 to Week 52/100. For each subject those endpoints are derived as the average of the changes from baseline to each post-baseline visit between Week 20 and Week 52/100, and between Week 28 and Week 52/100.
- Number and percentage of subjects with a gain in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline to each post-baseline visit

  Note: Subjects with BCVA value of 84 letters or more at a post-baseline visit will be considered as responders for the corresponding endpoint. This is to account for a ceiling effect, e.g. for the' ≥15-letter gain' endpoint, for those subjects with BCVA values at baseline ≥70 letters.
- Time to achieve gain in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline (or reaching a score of 84 or more)
- Number and percentage of subjects with a loss in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline to each post-baseline visit
- Number and percentage of subjects with an absolute BCVA ≥73 ETDRS letters at each post-baseline visit

Secondary efficacy endpoints related to dosing regimen:

- Proportion of subjects maintained at q12w up to Week 64 (after three q12w treatment intervals) and 100 (for brolucizumab treatment arms only). This endpoint will only be assessed in the year-2 analysis (Week 100), using the KM method as described for the key secondary endpoint at Week 52 (Section 2.7.2.4)
- Proportion of subjects maintained at q12w up to Week 64 (after three q12w treatment intervals), within those subjects that qualified for q12w at Week 36 (for brolucizumab treatment arms only). This endpoint will only be assessed in the year-2 analysis (Week 100), using the KM method as described for the key secondary endpoint at Week 52 (Section 2.7.2.4)
- Number and percent of subjects with q8w treatment need status assessed at Week 32
- Treatment status at Week 100. This endpoint will only be assessed in the year-2 analysis (Week 100)

Secondary efficacy endpoints related to anatomy:

- Change from baseline in Central Subfield Thickness (CSFT, as determined by SD-OCT from the central reading center) at each assessment visit

- Average change from baseline in CSFT over the period Week 40 through Week 52. For each subject this endpoint is derived as the average of the changes from baseline to Weeks 40, 44, 48, 52. Then the same will be derived over the period Week 88 through Week 100, considering the average of the changes from baseline to Weeks 88, 92, 96, 100. This endpoint will only be assessed in the year-2 analysis (Week 100)
- Average change from baseline in CSFT over the period Week 4 to Week 52/100
- Proportion of subjects with normal CSFT thickness (<280 microns) at each assessment visit
- Proportion of subjects with presence of SRF, IRF and simultaneous absence of SRF and IRF (i.e., presence of SRF and/or IRF) at each assessment visit
- Proportion of subjects with presence of leakage on FA at Weeks 52 and 100

Secondary efficacy endpoints related to the status of Diabetic Retinopathy (see Section 2.7.2.1):

- Proportion of subjects with a ≥2- and ≥3-step improvement or worsening from baseline in the ETDRS Diabetic Retinopathy Severity Scale (DRSS) score at each assessment visit
- Incidence of progression to proliferative diabetic retinopathy (PDR) as assessed by ETDRS-DRSS score ≥61 by Week 52 and Week 100, among non-PDR subjects at screening

2.7.2 Statistical hypothesis, model, and method of analysis

2.7.2.1 Confirmatory testing related to additional secondary efficacy endpoints

Confirmatory hypothesis testing for additional secondary endpoints will be performed in case the proof of non-inferiority related to BCVA is successful for the four hypotheses specified above (Section 2.5.2) for the primary and first key secondary endpoints (corresponding to H1, H2, H3 and H4 in Figure 2-1).

The additional hypotheses are linked to the endpoints below:

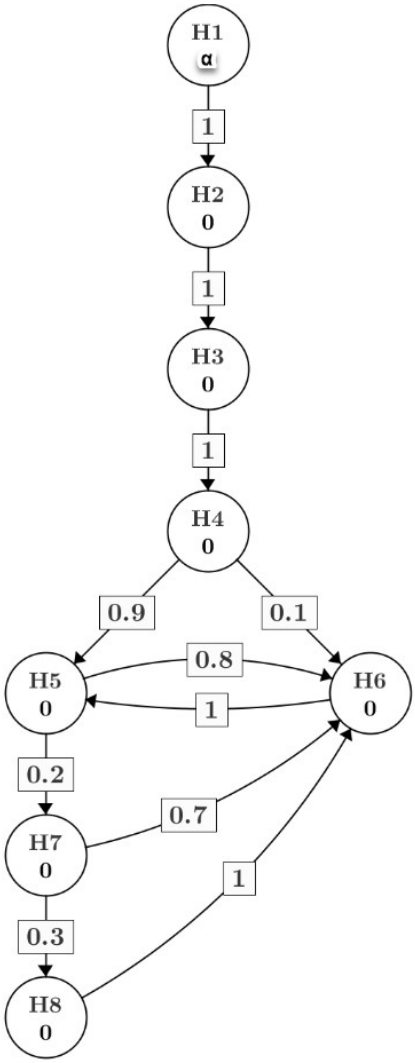
- H5. Average change from baseline in CSFT over the period Week 40 through Week 52 in the study eye
- H6. Absence of Fluid in the study eye at Week 52 (no= absence of SRF and IRF)
- H7. Change from baseline in CSFT at Week 4 in the study eye
- H8. Average change from baseline in BCVA over the period Week 40 through Week 52 in the study eye

All tests will be one-sided tests for superiority of brolucizumab 6 mg vs aflibercept 2 mg only (not brolucizumab 3 mg vs aflibercept 2 mg), i.e.,

- for greater reductions in the CSFT change from baseline (ANOVA as specified in Section 2.7.2.2) with brolucizumab compared to aflibercept
- for a higher proportion of subjects with absence of fluid (logistic regression as specified in Section 2.7.2.2) with brolucizumab compared to aflibercept

- for greater gains in the BCVA change from baseline (ANOVA as specified in Section 2.7.2.2) with brolucizumab compared to aflibercept

Figure 2-1 Multiple testing strategy



- Hypotheses H_1,..., H_8 are represented by circles with the initial significance levels. The arrow represents the direction in which the significance level is propagated throughout the graph and the number in the square box represents the proportion of the propagated significance level.

All the tests are performed at the level resulting from the graphical procedure. If a tested null hypothesis is rejected at the local significance level assigned to this null hypotheses, the alpha is passed on to other null hypotheses as per the graph.

As described in Section 2.5.2, the first four hypotheses will be tested sequentially in the order of their numbering (H_n, n=1, 2, 3, 4), i.e., confirmatory testing of the second, third or fourth hypotheses requires rejection of each preceding null hypothesis.

If each of the first four null hypotheses is rejected at a one-sided significance level of 0.025, the entire alpha will be distributed between the null hypotheses related to the superiority testing of

H_5 (90% of 0.025 = 0.0225), and H_6 (10% of 0.025 = 0.0025). This split is chosen by balancing out prior expectations about the study outcomes and the clinical importance of the endpoints.

The family-wise type I error rate will be controlled at the one-sided 2.5% level across the tested null hypotheses using the closed testing procedure specified by Figure 2-1 using the graphical method of Bretz, et al. (Bretz, et al 2009).

The basis for these tests for superiority will be the FAS with LOCF imputation/replacement of missing/censored data. For subjects who discontinue study treatment but continued in the study, data collected after the switch to alternative DME treatment in the study eye will be censored for the primary analysis.

2.7.2.2 General analysis specifications for secondary efficacy endpoints

All secondary efficacy endpoints listed in the above Section 2.7.1 will be summarized and presented descriptively, based on the FAS with LOCF imputation for missing or censored data if not otherwise specified. Details on data handling, such as missing values, are described in Section 2.7.3.

For year-1 CSR, the impact of COVID-19 pandemic on CSFT will be assessed with the subgroup analyses of exposed/non-exposed subjects and impacted/non-impacted subjects as defined in Section 2.5.4.3.

For year-2 CSR, the impact of COVID-19 pandemic on the the following secondary efficacy endpoints will be assesed with subgroup analysis of impacted and non-impacted subjects as defined in Section 2.5.4.3:

- Change from baseline in BCVA at each visit up to Week 100
- Proportion of subjects maintained at q12w up to 100 (for brolucizumab treatment arm only).
- Proportion of subjects maintained at q12w up to Week 100, within those subjects that qualified for q12w at Week 36 (for brolucizumab treatment arm only).
- Treatment status at Week 100.
- Change from baseline in Central Subfield Thickness (CSFT, as determined by SD-OCT from the central reading center) at each assessment visit
- Proportion of subjects with a ≥2-step improvement or worsening from baseline in the ETDRS Diabetic Retinopathy Severity Scale (DRSS) score at each assessment visit.

<u>Continuous endpoints:</u>

The continuous secondary endpoints related to BCVA and CSFT will be analyzed using ANOVA models. The estimates of least square means for each treatment and for the treatment differences brolucizumab – aflibercept, including 95% CIs for the treatment differences, will be presented.

For the ANOVA analysis of BCVA-related endpoints, baseline BCVA (≤65, >65 letters) and age category (<65, ≥65 years) will be considered as class variables. For the ANOVA analysis of CSFT, baseline CSFT (<450, ≥450-<650, ≥650 µm) will be used instead of baseline BCVA as a class variable.

The line plot on LSM (± SE) by visit will also be provided for all treatment arms.

Categorical variables:

For binary endpoints, frequency tables (count and percentage) will be provided by time point. In addition, proportions and treatment differences in proportions along with 95% CIs will be presented for each time point using a logistic regression with treatment, the corresponding baseline status (similar to the ones specified for the ANOVA models) and age categories as fixed effects.

Bar chart will be plotted by visit and treatment arm.

Time-to event variables:

Time-to-event variables such as the time to achieve gain in BCVA of ≥5 (respectively ≥10 and ≥15) letters from baseline (or reaching a score of 84 or more) will be analyzed using KM analysis. KM estimates on percent of subjects who achieve gain, together with 95% CI will be presented by visit. The median time (95% CI) to gain will also be constructed by treatment arm. KM curves presenting the cumulative probability of subjects with gain of ≥5 (respectively ≥10 and ≥15) letters from baseline will be provided by treatment arm.

2.7.2.3 ETDRS DRSS Score

Definition of Endpoints

The following endpoints related to diabetic retinopathy (DR) status will be analyzed:

- Subject status regarding a ≥2- and ≥3-step improvement or worsening from baseline in the ETDRS DRSS score at each assessment visit
- Incidence of progression to PDR as assessed by ETDRS-DRSS score of at least 61 by Week 52 and Week 100 (among non-PDR subjects at screening)

Those endpoints will be derived from the ETDRS-DRSS score assessed by the central reading center based on colour fundus (CF) photography images in the study eye at screening, Weeks 28, 52, 76 and exit/premature discontinuation visit.

When the ETDRS-DR severities are evaluable, they will be categorized using the following scores:

Table 2-3 Definition of DRSS: original scale

DRSS scale	Definition
10	DR absent
20	Microaneurysms only
35	Mild non-proliferative diabetic retinopathy (NPDR)
43	Moderate NPDR
47	Moderately severe NPDR
53	Severe NPDR
61	Mild PDR
65	Moderate PDR

DRSS scale	Definition
71	High-Risk PDR
75	Very high risk PDR
81	Advanced PDR
85	Very advanced PDR

Other recorded DRSS values (code 98: Indeterminable due to missing images, 99: Indeterminable due to upgradable images, 00: No images received) that are not related to an evaluable DR severity level will be handled as missing.

All DRSS values will be converted into a 12-point scale as defined in Table 2-4.

Table 2-4 Definition of DRSS: 12-point scale

12-point scale	Definition	Original DRSS
1	DR absent	10
2	Microaneurysms only	20
3	Mild NPDR	35
4	Moderate NPDR	43
5	Moderately severe NPDR	47
6	Severe NPDR	53
7	Mild PDR	61
8	Moderate PDR	65
9	High-Risk PDR	71
10	Very high-Risk PDR	75
11	Advanced PDR	81
12	Very advanced PDR	85

DR= diabetic retinopathy, DRSS= diabetic retinopathy severity score, NPDR= non-proliferative diabetic retinopathy, PDR= proliferative diabetic retinopathy.

Table 2-5 and Table 2-6 describe the definition of a 2-step and a 3-step change, respectively, for each (non-missing) baseline and post-baseline ETDRS based on the 12-point scale, as defined below:

- ≥2-step improvement: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≤-2
- ≥3-step improvement: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≤-3
- ≥2-step worsening: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≥2
- ≥3-step worsening: DRSS (12-point scale) at the visit – DRSS (12-point scale) at baseline ≥3

Table 2-5 Definition of 2-step change in DRSS on the 12-point scale

Baseline	Post-baseline		
	≥2-step improvement	No change or change <2 steps	≥2-step worsening
1	-	1, 2	3 or higher
2	-	1, 2 or 3	4 or higher
3	1	2, 3, or 4	5 or higher
4	1 or 2	3, 4, or 5	6 or higher
5	3 or lower	4, 5, or 6	7 or higher
6	4 or lower	5, 6, or 7	8 or higher
7	5 or lower	6, 7, or 8	9 or higher
8	6 or lower	7, 8, or 9	10 or higher
9	7 or lower	8, 9, or 10	11 or 12
10	8 or lower	9, 10, or 11	12
11	9 or lower	10, 11, or 12	-
12	10 or lower	11, 12	-

Table 2-6 Definition of 3-step change in DRSS on the 12-point scale

Baseline	Post-baseline		
	≥3-step improvement	No change or change <3 steps	≥3-step worsening
1	-	1, 2 or 3	4 or higher
2	-	1, 2, 3 or 4	5 or higher
3	-	1, 2, 3, 4 or 5	6 or higher
4	1	2, 3, 4, 5 or 6	7 or higher
5	1 or 2	3, 4, 5, 6 or 7	8 or higher
6	3 or lower	4, 5, 6, 7 or 8	9 or higher
7	4 or lower	5, 6, 7, 8 or 9	10 or higher
8	5 or lower	6, 7, 8, 9 or 10	11 or 12
9	6 or lower	7, 8, 9, 10 or 11	12
10	7 or lower	8 or higher	-
11	8 or lower	9 or higher	-
12	9 or lower	10 or higher	-

Analysis method

All DRSS analyses will be based on the 12-point scale shown in Table 2-4.

Proportions of subjects with ≥2- and ≥3-step improvement or worsening from baseline will be summarized using the FAS by assessment visit. Bar chart will be plotted by assessment visit and treatment arm.

For the proportions of subjects with ≥2-step change from baseline at Week 52 (and similarly for ≥3-step change), the 95% confidence intervals (CIs) for the proportions in all treatment arms, the differences in proportions between brolucizumab and aflibercept treatment arms and the 95% CI for the difference will be calculated using a logistic regression with treatment, the

corresponding baseline DRSS (score ≤43, ≥47 from the original scale or ≤4, ≥5 from the 12-level scale) and age category (<65, ≥65 years) as fixed effects.

The impact of COVID-19 pandemic on the proportion of subjects with ≥2-step improvement or worsening from baseline will be assessed with the subgroup analyses of exposed/non-exposed subjects and impacted/non-impacted subjects as defined in Section 2.5.4.3.

The proportion of subjects who progress to PDR, as assessed by ETDRS-DRSS score ≥61 by Week 52 and Week 100, will be summarized among the subset of non-PDR subjects at screening (ETDRS-DRSS score <61 at screening).

2.7.2.4 Analysis of year-2 dosing regimen endpoints

The estimate for the proportion of patients with a positive q12w treatment status will be derived from Kaplan-Meier (KM) time-to-event analyses for the event 'first q8w-need' as described for the key secondary endpoint at Week 52 in Section 2.6.2. Specifically, the proportion of subjects with a positive q12w treatment status will be derived as follows requiring 'sufficient duration of effect' (as assessed by q8w-need) together with 'sufficient efficacy and safety':

- For the 'sufficient duration of effect' requirement subjects will need to have the status of 'q8w-need =no' at Weeks 32, 36, 48, 60, 72 and all scheduled q12w treatment visits unless the 'q8w-need = yes' is confounded by reasons other than lack of efficacy and/or safety (see censoring details below)
- The requirement regarding 'sufficient efficacy and safety' will be addressed by considering subjects – even without an explicit 'q8w-need = yes' – as having a negative q12w status in case any of the following confounding factors is attributable to lack of efficacy and/or lack of safety of the study treatment: early treatment/study discontinuation, missed DAA. The q8w-need assessment is imputed as "Yes" at the valid DAA visit (DAA visit with the option to change to q8w) following early treatment/study discontinuation due to lack of efficacy and/or lack of safety of the study treatment (applies to both missing and non-missing DAAs).

Intercurrent events associated with missing or confounded data regarding the q12w treatment status that are attributable to reasons other than lack of efficacy and/or safety will be handled the same as will be done for the key secondary endpoint.

While for the analysis of the overall q12w proportion all subjects in the FAS will be considered, the analysis of the proportion of subjects maintained at q12w up to Week 100, within those subjects that qualified for q12w at Week 36, will be based on the subset of FAS subjects with no identified q8w-need at Week 32 and Week 36.

Censoring rules related to the q12w treatment status analysis as described in Section 5.6 apply to the q12w analysis for year-2 CSR. Subjects will be censored at Week 100 if they completed Week 100 without any PD/AR listed in Table 5-6 for DAA.

Treatment status at Week 100 will be summarized based on subjects who completed the study treatment.

2.7.3 Handling of missing values/censoring/discontinuations

Missing data for all the secondary efficacy endpoints will be imputed using the LOCF method unless specified otherwise.

For the LOCF method, missing data will be imputed by the value of the last available non-missing post-baseline observation. For subjects who discontinue treatment but continue in the study, data collected after the start of alternative DME treatment in the study eye will be censored for the analysis. Censored data will be replaced by the last available observation prior to the start of alternative DME treatment in the study eye.

Missing baseline values will not be imputed. For subjects with no post-baseline values (scheduled or unscheduled), the baseline value will be carried forward, as a conservative approach.

For endpoints related to presence of SRF and/or IRF, if baseline visit is reported as ”Cannot Grade”, then it will be considered as ”Absent”; if post-baseline visit is reported as ”Cannot Grade”, then it will be considered as missing and LOCF method for imputation will be applied.

For the presence of leakage on FA, if baseline visit is reported as ”Cannot Grade”, then it will be considered as missing; if post-baseline visit is reported as ”Cannot Grade”, then it will be considered as missing and LOCF method for imputation will be applied.

2.8 Safety analyses

Safety endpoints are based on the variables from safety assessments, which include:

- Extent of exposure (see Section 2.4.1)
- Adverse events
- Ophthalmic examinations
- Vital signs
- Laboratory results
- Imaging parameters

There are no formal safety hypotheses in this study. All safety analyses will be performed using the Safety Analysis Set.

Except for imputation of partial dates for AEs, no imputations will be performed for missing values in the safety analyses.

In all summary tables, unless otherwise specified (e.g. for AE tables), data collected after the subject discontinued study treatment and started alternative DME treatment in the study eye will be censored (data on the day the subject started alternative DME treatment will be included).

2.8.1 Adverse events (AEs)

A treatment-emergent adverse event (TEAE) is defined as any adverse event that develops after exposure to the study treatment or any event already present that worsens following exposure

to the study treatment. Only treatment-emergent adverse events will be presented in the summary tables.

Adverse events will be coded using the MedDRA dictionary and presented by system organ class (SOC), preferred term (PT) and treatment arm. Treatment-emergent AEs will be analyzed based on the number and percentage of subjects with at least one AE in the category of interest.

The number (and proportion) of subjects with TEAEs will be summarized at each analysis timepoint (Week 52, Week 100) in the following ways:

Table 2-7 TEAE summary

TEAE summary	AE categories		
	Ocular AE in the study eye	**Ocular AE in the fellow eye**	**Non-ocular AE**
AEs by primary SOC and PT	Y#		Y#
AEs by primary SOC and PT (including events with onset date after start of alternative DME treatment)	Y	Y	Y
Frequent AEs by PT†	Y		Y
AEs by maximum severity, SOC and PT	Y		Y
AEs related to study treatment by SOC and PT	Y		Y
AEs related to injection procedure by SOC and PT	Y		
AEs leading to permanent discontinuation of study treatment by SOC and PT	Y		Y
AEs leading to temporary interruption of study treatment by SOC and PT	Y		Y
SAEs by SOC and PT	Y#		Y#
SAEs by SOC and PT (including events with onset date after start of alternative DME treatment)	Y	Y	Y
SAEs related to study treatment by SOC and PT	Y		Y
SAEs related to injection procedure by SOC and PT	Y		

†≥2 % (or other cutting point as appropriate) in any treatment group for a given PT.
including separate summary tables in year-1 CSR for exposed/non-exposed and impacted/non-impacted subjects to COVID-19 as defined in Section 2.5.4.3, and including separate summary table in year-2 CSR for impacted and non-impacted subjects to COVID-19 as defined in Section 2.5.4.3

In all summary tables listed above, unless otherwise specified, data collected after the subject discontinued study treatment and started alternative DME treatment in the study eye will be censored.

If an AE started on the same day as the start of alternative DME treatment for a subject, the AE will be excluded from the summary table, unless this AE led to study drug withdrawal (in such a case, the AE would be included in the analysis).

Subject listings of all adverse events will be provided. Deaths and SAEs (i.e., other serious or clinically significant non-fatal adverse events) will be listed separately.

The MedDRA version used for reporting the AEs will be described in a footnote.

2.8.1.1 Adverse events of special interest / grouping of AEs

Incidence of adverse events of special interest (AESI) will be tabulated by treatment arm.

AESIs and other safety topics of interest will be identified via the RTH258 electronic case retrieval strategy (eCRS). The eCRS that is current at the time the database lock will be used and AESIs and other safety topics of interest will be identified where the flag Core Safety Topic Risk (SP) = 'Y'.

2.8.1.2 Adverse event reporting for clinical trial safety disclosure

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on TEAEs which are not serious adverse events with an incidence greater than 5% and on TEAEs and SAEs suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population. Ocular TEAEs for study eye and fellow eye will be considered separately.

If for the same subject, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is >1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE (respectively non-SAE) has to be checked in a block e.g., among AEs in a ≤1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment, and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.8.2 Deaths

A summary of treatment emergent deaths will be presented by primary SOC and PT.

All deaths recorded in the clinical database will be listed.

2.8.3 Laboratory data

Laboratory data will be presented graphically using boxplots of absolute change from baseline values by treatment arm and visit. No summary by visit tables will be provided.

A summary table with counts and percentage of subjects satisfying the criteria representing clinically relevant abnormalities given in Section 5.3 at any visit will be presented. A listing for subjects satisfying at least one criterion in Table 5-1 at any visit will also be presented.

2.8.4 Other safety data

2.8.4.1 Ophthalmic examinations

Descriptive summaries of pre-injection change from baseline in intraocular pressure (IOP) values for the study eye will be presented graphically at each study visit by treatment arm, considering line plots of the mean change in IOP values with error bars representing ± SE. The x-axis will be study visit and the y-axis will be the change from Baseline value. No summary by visit tables will be provided.

The number and percentage of subjects with pre-injection IOP >30 mmHg at any visit will be summarized.

Post-injection IOP is to be assessed approximately within 60 minutes after injection and if ≥25 mmHg, the assessment should be repeated until back to normal. Summary tables with counts and percentage of subjects with an IOP increase of ≥10, ≥20 mmHg from pre-injection to post-injection at any visit for the study eye will be presented.

A summary table with counts and percentage of subjects with observed pre-injection IOP ≥ 21 mmHg at 3 consecutive scheduled visits will be presented.

A visit with missing pre-injection IOP is considered to meet the ≥21 mmHg criterion if the preceeding and the following visits meet the criterion that pre-injection IOP ≥21 mmHg. For example, if schedule visit X has missing pre-injection IOP and pre-injection IOP ≥21 mmHg is observed for both visit X-1 and X+1, the subject is considered to meet the criteria at visit X as well.

A listing for subjects with any post-injection IOP increase of ≥10 mmHg from pre-injection IOP and a listing of subjects with any IOP >30 mmHg will be presented.

The abnormal findings via slit-lamp and indirect fundus examinations deemed as clinically significant by the investigator and reported as AE/SAE will be included in the safety analysis on AE/SAE.

2.8.4.2 Loss in BCVA

The number and percentage of subjects with a loss in BCVA ≥15, ≥30 letters (study eye) from baseline to each visit, to the last visit, and maximum loss at any visit will be presented.

BCVA data (study eye) for subjects presenting loss in BCVA ≥15 letters from baseline at any post-baseline visit will be listed.

2.8.4.3 Vital signs

A summary table with counts and percentage of subjects satisfying the criteria given in Table 5-2 of the Section 5.3 at least one visit will be presented. A listing for subjects satisfying at least one criterion in Table 5-2 will also be presented.

A line plot of mean change from baseline in the vital sign parameter by study visit and treatment arm with error bars representing ±1 standard error will be presented. The x-axis will be study visit and the y-axis will be the mean change from baseline value.

2.8.4.4 Imaging parameters

Pre-defined imaging parameters in the study eye typically associated with intraocular inflammation and/or retinal vascular occlusion as assessed by the CRC will be listed per visit. No summary table will be provided.

2.9 Pharmacokinetic endpoints

Not Applicable.

2.10 Anti-drug antibodies

Collection of blood sample for ADA assessment for brolucizumab will be done at Screening, Weeks 4, 12, 24, 36, 52, and 76 prior to the injection/sham, and at exit/premature discontinuation.

[redacted]

[redacted]

Systemic exposure of brolucizumab will be measured concomitantly with ADA levels for interpretation purposes, no pharmacokinetic parameters will be determined from brolucizumab systemic exposure. Systemic exposure data will be summarized and listed.

2.11 Patient-reported outcomes

The Visual Function Questionnaires (VFQ-25) will be scored (total and subscale scores) at Baseline and Weeks 28, 52, 76 and 100 visits. Absolute scores and the absolute changes from baseline will be calculated and summarized descriptively using the FAS.

Further details on the scoring algorithm and analysis are provided below.

Each subscale score has a range of 0 to 100 inclusive and will be calculated from the re-scaled raw data as described in Table 2-8. A missing response will not be re-scaled (except for the response to question 15c, see below, which will be re-set to 0 if the response to question 15b is 1).

The answers to questions will be re-scaled as follows to calculate the total and subscale scores.

Table 2-8 Rescaling of VFQ-25 questions

Answer to question	Rescaling for questions 1, 3, 4 and 15c	Rescaling for question 2	Rescaling for questions 5-14, 16 and 16a	Rescaling for questions 17-25
1	100	100	100	0
2	75	80	75	25
3	50	60	50	50
4	25	40	25	75
5	0	20	0	100
6	N/A	0	N/A*	N/A

Note: * 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". Subscales will be calculated where at least one of the (re-scaled) questions contributing to that subscale is non-missing, and otherwise set to missing.

- Note that the answer to question 15c will subsequently be adjusted based on the answer to question 15b.
    - If the answer to 15b is 1 then the answer to 15c will be re-set to 0.
    - If the answer to 15b is 2 or 3 then the answer to 15c will be re-set to missing

The general health rating is the re-scaled answer to question 1.

The scales and corresponding questions are shown in Table 2-9.

Table 2-9 Questions contributing to VFQ subscales

Subscale	Questions
General vision	2
Ocular pain	4 and 19

Near activities	5, 6 and 7
Distance activities	8, 9 and 14
Social functioning	11 and 13
Mental health	3, 21, 22 and 25
Role difficulties	17 and 18
Dependency	20, 23 and 24
Driving	15c, 16 and 16a
Color vision	12
Peripheral vision	10

The composite score is the average of the 11 subscales shown in Table 2-9. It will be set to missing if at least six of the subscales are missing.

Descriptive summary statistics for change from baseline to post baseline VFQ assessments will be presented using the FAS for the composite and subscale scores. Mean changes from baseline to each post baseline VFQ assessments visits will be compared between the brolucizumab arms and the aflibercept arm. Appropriate statistical methods (e.g. pairwise ANCOVA model with treatment as a fixed effect factor and corresponding baseline value of the endpoint in the model) will be used for treatment arm comparisons. Additionally, descriptive statistics will also be presented for the general health score. All analyses will be performed on the subscales values.

The VFQ-25 composite score and subscale scores will not be listed.

[redacted]

2.13 Interim analysis

The analysis based on the Week 52 data will be the primary efficacy and safety analysis for this study. The database includes all data up to Week 52 from when all randomized subjects have completed the Week 52 visit or terminated the study prior to (or at) Week 52.

The results of this analysis will be reported in the year-1 CSR only and will include the analysis of the overall population and subgroup analyses Japan vs non-Japan for the following data:

- Subject disposition (Section 2.3.1)
- Demographics and baseline characteristics (Section 2.3.2)
- Study treatment exposure (Section 2.4.1)
- Subgroup analysis of the primary and key secondary efficacy variables (Section 2.2.1)
- Change from baseline in BCVA at each visit up to Week 52 (Section 2.7)
- Change from baseline in CSFT at each visit up to Week 52 (Section 2.7)
- Proportion of subjects with presence of SRF, IRF and simultaneous absence of SRF and IRF (i.e., presence of SRF and/or IRF) at each assessment visit (Section 2.7)
- Proportion of subjects with a ≥2- and ≥3-step improvement or worsening from baseline in the DRSS score at each assessment visit (Section 2.7)
- The number (and proportion) of subjects with ocular AEs/SAEs in the study eye and non-ocular AEs/SAEs, up to Week 52 (Section 2.8.1)

Japan subgroup analyses may be performed in the year-2 CSR for some efficacy and safety endpoints.

A second planned interim analysis may be performed by locking the Week 76 data in case of regulatory request of supplemental data to be submitted during the review period.

3 Sample size calculation

A sample size of 160 subjects per arm will allow demonstration of non-inferiority (NIM of 4 ETDRS letters) of brolucizumab 6 mg or 3 mg vs. aflibercept 2 mg with respect to the BCVA

change from baseline at Week 52, with 90% power (disregarding the dependence within the sequential testing procedure, i.e. local power for 3 mg) at a one-sided alpha level of 0.025, assuming equal means and a common standard deviation of 11 letters. Assuming that averaging over the 4 time points will not lead to an increase in the standard deviation a power of at least 90% can also be expected for its corresponding non-inferiority claim.

To account for a drop-out rate of 10%, a total of 534 (178 per arm) subjects will need to be randomized.

4 Change to protocol specified analyses

There is no change to the protocol specified analyses in terms of endpoints.

Confirmatory hypothesis testing in relation to additional secondary endpoints is introduced in Section 2.7.2.1. Some changes compared to the protocol specified analyses are considered in the current statistical analysis plan before database lock in order to implement the Novartis internal process on SAP simplification (LEAN):

Protocol section	Protocol wording	Change in the SAP
9.2	Demographics and baseline characteristics will be summarized with descriptive statistics for all analysis sets by treatment group and overall.	Demographics and baseline characteristics will be summarized with descriptive statistics for the FAS by treatment arm and overall.
9.3	Descriptive statistics for exposure to study treatment will be provided for the safety set, FAS and PPS	Descriptive statistics for exposure to study treatment will be provided for the Safety set.
[redacted]		
9.5.2	Laboratory data and vital signs will be summarized by presenting shift tables using extended normal ranges (as provided by the central laboratory) with thresholds representing clinical relevant abnormality and by presenting descriptive statistics of raw data and change from baseline. Values outside the extended normal range will be listed by subject and treatment arm and flagged in data listings.	No summary by visit tables will be provided. A summary table with counts and percentage of subjects satisfying the criteria representing clinically relevant abnormalities at any visit will be presented. A listing for subjects satisfying at least one criterion in Table 5-1 at any visit will also be presented.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

No imputation will be made to the start date and end date of study treatment.

5.1.2 AE date imputation

5.1.2.1 AE start date imputation

The following table explains the notation used in the logic matrix below. Please note that completely missing start dates will not be imputed.

	Day	Month	Year
Partial Adverse Event Start Date	Not used	MON	YYYY
Treatment Start Date	Not used	TRTM	TRTY

The following matrix explains the logic behind the imputation.

	MON MISSING	MON < TRTM	MON = TRTM	MON > TRTM
YYYY MISSING	(1) No convention	(1) No convention	(1) No convention	(1) No convention
YYYY < TRTY	(2.a) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start
YYYY = TRTY	(4.a) Uncertain	(4.b) Before Treatment Start	(4.c) Uncertain	(4.c) After Treatment Start
YYYY > TRTY	(3.a) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start

Before imputing AE start date, find the AE start reference date.

1. If the (imputed) AE end date is complete and the (imputed) AE end date < treatment start date then AE start reference date = min (informed consent date, earliest visit date).
2. Else AE start reference date = treatment start date.

Impute AE start date -

1. If the AE start date year value is missing, the date uncertainty is too high to impute a rational date. Therefore, if the AE year value is missing, the imputed AE start date is set to NULL.
2. If the AE start date year value is less than the treatment start date year value, the AE started before treatment. Therefore:
   a. If AE month is missing, the imputed AE start date is set to the mid-year point (01JulYYYY).
   b. Else if AE month is not missing, the imputed AE start date is set to the mid-month point (15MONYYYY).
3. If the AE start date year value is greater than the treatment start date year value, the AE started after treatment. Therefore:

a. If the AE month is missing, the imputed AE start date is set to the year start point (01JanYYYY).

b. Else if the AE month is not missing, the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).

4. If the AE start date year value is equal to the treatment start date year value:

a. And the AE month is missing the imputed AE start date is set to the AE reference start date + 1 day.

b. Else if the AE month is less than the treatment start month, the imputed AE start date is set to the mid-month point (15MONYYYY).

c. Else if the AE month is equal to the treatment start date month or greater than the treatment start date month, the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).

If complete (imputed) AE end date is available and the imputed AE start date is greater than the (imputed) AE end date, then imputed AE start date should be set to the (imputed) AE end date.

5.1.2.2 AE end date imputation

1. If the AE end date month is missing, the imputed end date should be set to the earliest of the (31DECYYYY, date of death).
2. If the AE end date day is missing, the imputed end date should be set to the earliest of the (last day of the month, date of death).
3. If AE year is missing or AE is ongoing, the end date will not be imputed.
4. If the imputed AE end date is less than the existing AE start date then use AE start date as AE end date.

5.1.3 Concomitant medication date imputation

5.1.3.1 Concomitant medication start date

In order to classify a medication as prior and prior/concomitant, it may be necessary to impute the start date.

Completely missing start dates will be set to one day prior to treatment start date. As a conservative approach, such treatments will be classified as prior and concomitant (and summarized for each output).

Concomitant treatments with partial start dates will have the date or dates imputed.

The following table explains the notation used in the logic matrix

	Day	**Month**	**Year**
Partial CMD Start Date	Not used	MON	YYYY
Treatment Start Date	Not used	TRTM	TRTY

The following matrix explains the logic behind the imputation.

	MON MISSING	MON < TRTM	MON = TRTM	MON > TRTM
YYYY MISSING	(1) Uncertain	(1) Uncertain	(1) Uncertain	(1) Uncertain
YYYY < TRTY	(2.a) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start	(2.b) Before Treatment Start
YYYY = TRTY	(4.a) Uncertain	(4.b) Before Treatment Start	(4.a) Uncertain	(4.c) After Treatment Start
YYYY > TRTY	(3.a) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start	(3.b) After Treatment Start

1. If the CM start date year value is missing, the imputed CM start date is set to one day prior to treatment start date.
2. If the CM start date year value is less than the treatment start date year value, the CM started before treatment. Therefore:
   a. If the CM month is missing, the imputed CM start date is set to the mid-year point (01JulYYYY).
   b. Else if the CM month is not missing, the imputed CM start date is set to the mid-month point (15MONYYYY).
3. If the CM start date year value is greater than the treatment start date year value, the CM started after treatment. Therefore:
   a. If the CM month is missing, the imputed CM start date is set to the year start point (01JanYYYY).
   b. Else if the CM month is not missing, the imputed CM start date is set to the month start point (01MONYYYY).
4. If the CM start date year value is equal to the treatment start date year value:
   a. And the CM month is missing or the CM month is equal to the treatment start date month, then the imputed CM start date is set to one day prior to treatment start date.
   b. Else if the CM month is less than the treatment start date month, the imputed CM start date is set to the mid-month point (15MONYYYY).
   c. Else if the CM month is greater than the treatment start date month, the imputed CM start date is set to the month start point (01MONYYYY).

If complete (imputed) CM end date is available and the imputed CM start date is greater than the (imputed) CM end date, then imputed CM start date should be set to the (imputed) CM end date.

5.1.3.2 Concomitant medication (CM) end date imputation

1. If the CM end date year value is missing, the date uncertainty is too high to impute a rational date. Therefore, if the CM end year value is missing or ongoing, the imputed CM end date is set to NULL.

2. Else, if the CM end date month is missing, the imputed end date should be set to the earliest of the (treatment end date, 31DECYYYY, date of death).

3. If the CM end date day is missing, the imputed end date should be set to the earliest of the (treatment end date, last day of the month, date of death).

4. If the imputed CM end date is less than the existing CM start date, use the CM start date as the imputed CM end date.

5.1.4 Medical history date of diagnosis imputation

Completely missing dates and partially missing end dates will not be imputed. Partial dates of diagnosis will be compared to the treatment start date.

1. If DIAG year < treatment start date year
    a. and DIAG month is missing, the imputed DIAG date is set to the mid-year point (01JULYYYY).
2. else if DIAG month is not missing, the imputed DIAG date is set to the mid-month point (15MONYYYY).
3. If DIAG year = treatment start date year
    a. and (DIAG month is missing OR DIAG month is equal to treatment start month), the imputed DIAG date is set to one day before treatment start date.
    b. else if DIAG month < treatment start month, the imputed DIAG date is set to the midmonth point (15MON YYYY).
    c. else if DIAG month > treatment start month => data error.
4. If DIAG year > treatment start date year => data error.

5.2 AEs coding/severity

AEs are coded using the MedDRA terminology.

AEs severity is assessed by investigators according to the following:

- mild: usually transient in nature and generally not interfering with normal activities
- moderate: sufficiently discomforting to interfere with normal activities
- severe: prevents normal activities

5.3 Laboratory parameters and vital signs derivations

Table 5-1 Clinically notable laboratory values

Test	Conventional Units	Critical Low	Critical High	Standard Units	Critical Low	Critical High	Non-nemeric
Calcium	mg/dL	< 6.0	> 13.0	mmol/L	< 1.50	> 3.25	
Creatinine		NA	>3xULN				
Glucose	mg/dL	< 40	> 450	mmol/L	< 2.2	> 25.0	
Potassium	mEq/L	< 2.8	> 6.2	mmol/L	< 2.8	> 6.2	
Sodium	mEq/L	< 120	> 160	mmol/L	< 120	> 160	

HCG							Negative, inconclusive
Hematocrit	%	< 20	> 60	V/V	< 0.20	> 0.60	
Hemaglobin	g/dL	< 6.0	> 20.0	g/L	< 60	> 200	
Platelet	X10E3/uL	< 50	> 999	X10E9/L	< 50	> 999	
WBC	X10E3/uL	< 2.0	> 35.0	X10E9/L	< 2.0	> 35.0	

Table 5-2 Clinically notable vital signs

Variable	Category	Critical values
Systolic blood pressure (mmHg)	High	Either >180 with an increase from baseline >30 or >200 absolute
	Low	Either <90 with a decrease from baseline >30 or <75 absolute
Diastolic blood pressure (mmHg)	High	Either >105 with an increase from baseline >20 or >115 absolute
	Low	Either <50 with a decrease from baseline > 20 or <40 absolute
Pulse rate (bpm)	High	Either >120 with an increase from baseline of >25 or > 130 absolute
	Low	Either <50 with a decrease from baseline >30 or <40 absolute

5.4 Statistical models

5.4.1 Primary and first key secondary analysis

The primary endpoint (change from baseline in BCVA at Week 52) and first key secondary endpoint (average change from baseline in BCVA over the period Week 40 through Week 52) will be analyzed using ANOVA models.

The ANOVA models will be fitted separately for brolucizumab 3mg vs. Aflibercept 2mg and for brolucizumab 6mg vs. Aflibercept 2mg.

Analysis of Variance (ANOVA)
The following ANOVA model will be used for the primary and first key secondary efficacy endpoints:
<change from Baseline in BCVA at Week 52> <average change from Baseline in BCVA from Week 40 to Week 52> = intercept + treatment + Baseline BCVA category + age category + error.
For the above analysis, the data structure is one record per subject. The SAS Proc MIXED will be used to perform the ANOVA analyses.

Mixed Model Repeated Measures (MMRM)
The following MMRM model will be used for the supportive analysis of the primary and first key secondary efficacy variables:
*<change from Baseline in BCVA at Week 52> <average change from Baseline in BCVA from Week 40 to Week 52> = intercept + treatment + Baseline BCVA category + age category + visit + treatment*visit + error.*

The SAS Proc MIXED will be used to perform the MMRM analyses.

Note: For the above MMRM analysis, the data structure is one record per FAS subject per scheduled visit. The data will include all subjects and have records for all scheduled visits, regardless of whether the assessment was missed or not at a given visit. Missing values will NOT be imputed using LOCF. Instead, the value will be passed to the model as missing.

5.4.2 Other secondary efficacy analysis

5.4.2.1 ANCOVA model for continuous variables

The continuous efficacy variables (such as VQF-25 score change from baseline) will be analyzed using an ANCOVA model adjusted for treatment, age category, and the corresponding baseline VQF-25 score.
The SAS Proc MIXED will be used to perform the ANCOVA analyses

5.4.2.2 Logistic regression for proportion variables

The binary efficacy variables will be analyzed using the logistic regression model adjusted for treatment, age category, corresponding baseline variables, and other covariates if necessary, using the FAS.

The SAS Proc LOGISTIC will be used.
Note:

- For the above analyses, the data structure is one record per subject and visit.The least square mean estimates obtained from the above model are for the log-odds ratios.
- The estimated difference in proportions and the corresponding 95% confidence intervals will be obtained by applying the bootstrap method. The pseudo SAS code to derive the treatment difference and 95% CI from the least square mean output of the fitted model will be provided in the programming specification document.

5.4.2.3 KM estimate for time to event variables

Within the brolucizumab treatment arms, the proportion of subjects maintained at q12w up to Week 52 will be estimated from Kaplan Meier time-to-event analyses for the event ‘first q8w-need’, applying event allocations (in case of lack of efficacy and/or lack of safety) and censoring as described in Section 2.6.2.

A corresponding 95% CI will be derived from the LOGLOG transformation, using SAS Proc Lifetest, with CONFTYPE = LOGLOG.

5.5 Rule of exclusion criteria of analysis sets

Important protocol deviations are defined in the Protocol Deviations Requirements Document. Table 5-3 includes the important protocol deviations which lead to exclusion of a subject from one or more analysis sets for the Week 52 analysis:

Table 5-3 Important protocol deviations leading to exclusion from analyses

Deviation ID	Description of Deviation	Exclusion in Analyses
M_INCL01_ICF not obtained	Written informed consent not obtained	Exclude from all analyses
P_INCL02_Age less than 18 yrs	Patient less than 18 years of age at baseline	Exclude from PP analysis
M_INCL03_Diabetes eligibility criteria	Patients without diabetes mellitus or HbA1c of more than 10% at screening or insufficient diabetes management at screening or baseline	Exclude from PP analysis
P_INCL04_No visual impairment (study eye)	Study Eye: no visual impairment due to DME as per BCVA or CSFT criteria	Exclude from PP analysis
M_EXCL01_Confounding condition in study eye	Study Eye: Confounding ocular concomitant conditions or ocular disorders	Exclude from PP analysis
M_EXCL02_Confounding concomitant medications or procedures in study eye	Study Eye: Confounding concomitant medications or procedures	Exclude from PP analysis
M_TRT01_Wrong IP administered	Wrong IP administered during the study	Exclude from PP analysis, unless brolucizumab 3mg was given instead of brolucizumab 6mg, or brolucizumab 6mg was given instead of brolucizumab 3mg
M_TRT02_Under-treatment during loading phase	Under-treatment during loading phase; missed active treatment (not due to any safety event)	Exclude from PP analysis
M_TRT03_Over treatment	Over treatment, received active when schedule was for sham /no treatment	Exclude from PP analysis
M_TRT04_Under-treatment after loading phase	Under-treatment after loading phase; missed active treatment (not due to any safety event)	Exclude from PP analysis if any missed active between W40 and W48 inclusive, or if at least 2 missed consecutive active doses (not due to safety); Otherwise include in all analyses
M_OTH01_Masking process not followed	Masking process not followed as per protocol with impact on data integrity	Exclude from PP analysis
M_OTH02_Any other PD	Any other protocol deviation with impact on the efficacy assessments or safety of the patient	Exclude from PP analysis
P_WITH01_Treatment but consent withdrawn	Subject withdrew consent but continued to receive study medication	Exclude from PP analysis

Table 5-4 lists the non-protocol deviations (analysis restrictions) that may lead to exclusion from per-protocol analysis. Analysis restrictions (ARs) address limitations in the evaluability which result from missing or confounded data with underlying background not qualifying as a PD (e.g. early study terminations, early treatment discontinuations, missing DAA or BCVA assessments).

Subject evaluability is based on two components:

- Exclusion from an analysis set
- Censoring of specific data points from an analysis (see Section 5.6).

The consequence of an AR on the evaluability depends on the underlying reason, while three different categories of reason are considered:

- Lack of efficacy of the study treatment (=1)
- Lack of safety / tolerability of the study treatment (=2)
- Other (=0)

Remark: Based on the concept of PD's, their underlying reason will always be '0'.

As a general rule, ARs with a reason of 1 or 2 do not lead to an exclusion from any analysis set, as a potential link between exclusion reason and treatment would constitute a source for systematic bias.

Rules of determination of ARs by programming will be specified in the Programming Data Specifications (PDS) documentation.

Table 5-4 Non-protocol deviations (analysis restrictions)

AR ID	Description of AR	Category of reason	Exclusion in Analyses
AR_EST_01	Early study termination due to lack of efficacy	1	Include in all analyses
AR_EST_02	Early study termination due to lack of safety	2	Include in all analyses
AR_EST_03	Early study termination due to reasons other than lack of efficacy/safety	0	Exclude from PP analysis if before Week 40 Otherwise include in all analyses
AR_ETD_01	Early study treatment termination due to lack of efficacy	1	Include in all analyses
AR_ETD_02	Early study treatment termination due to lack of safety	2	Include in all analyses
AR_ETD_03	Early study treatment termination due to reasons other than lack of efficacy/safety	0	Exclude from PP analysis if before Week 40 Otherwise include in all analyses
AR_MD_01	No valid BCVA assessment between Week 40 and Week 52	0	Exclude from PP analysis
AR_MD_02	Missing DAA due to lack of safety	2	Include in all analyses
AR_MD_03	Missing DAA due to reasons other than lack of safety	0	Include in all analyses

Table 5-5 describes subject classification with regards to analysis sets:

Table 5-5 Subject classification

Analysis Set	PD ID that may cause subjects to be excluded	Non-PD (AR) ID that cause subjects to be excluded
RAN	M_INCL01_ICF not obtained	Not Randomized;
FAS	M_INCL01_ICF not obtained	Not in the RAN; Did not receive at least one study injection
SAF	M_INCL01_ICF not obtained	Did not receive at least one study injection
PPS	M_INCL01_ICF not obtained P_INCL02_Age less than 18 yrs, M_INCL03_Diabetes eligibility criteria, P_INCL04_No visual impairment (study eye), M_EXCL01_Confounding condition in study eye,	Not in the FAS AR_EST_03, AR_ETD_03, AR_MD_01

Analysis Set	PD ID that may cause subjects to be excluded	Non-PD (AR) ID that cause subjects to be excluded
	M_EXCL02_Confounding concomitant medications or procedures in study eye, M_TRT01_Wrong IP administered, M_TRT02_Under-treatment during loading phase, M_TRT03_Over treatment, M_TRT04_Under-treatment after loading phase, M_OTH01_Masking process not followed, M_OTH02_Any other PD, P_WITH01_Treatment but consent withdrawn	

5.6 Censoring rules for analysis

Protocol deviations (PDs) and analysis restrictions (ARs) that are considered to be critical for the subject evaluability regarding the primary and key secondary endpoints are described in Section 5.5.

The focus of the ARs is the identification of censoring related to the analysis of BCVA and q12w proportion as derived from DAA and described in Section 2.6.2. Censoring for DAA is only applied in case the underlying reason for a confounded DAA is assessed as '0'. Censoring of BCVA and DAA applies both to the year-1 analysis and year-2 analysis for the FAS, and only to the year-1 analysis for the PPS.

Table 5-6 summarizes the concepts of censoring for the key parameters BCVA and q12w-status/DAA applied to the two efficacy analysis sets, FAS and PPS, as well as the details for the timing of censoring for BCVA and DAA.

In case a subject has multiple PDs/ARs with impact on subject's evaluability the following rules are applied:

- A subject is excluded from an analysis set if at least one PD or AR with this consequence was identified (see Table 5-5). This rule is built on the concept of the medical assessment of the 'reason' which considers the reason of an earlier event to potentially also be the reason for following PDs or ARs.
- In case of multiple censoring time points censoring will be performed at the earliest.

Table 5-6 Censoring concepts for BCVA and DAA

Analysis Set	Censoring concept for BCVA	Censoring concept for DAA
FAS	Censoring of BCVA data after switch to alternative DME treatment in the study eye: imputation using the last observation collected prior to the start of alternative DME treatment (see section 2.5.3) No other censoring related to PDs or ARs.	M_TRT01_Wrong_IP_administered: censoring at the last valid DAA visit at or prior to the PD visit M_TRT02_Under-treatment during loading phase: censoring at baseline M_TRT03_Over treatment: censoring at the last valid DAA visit at or prior to the PD visit M_TRT04_Under-treatment after loading phase: censoring at the last valid DAA visit at or prior to the PD visit M_COMD01_Prohibited medication or procedure: censoring at the last valid DAA prior to the start of the prohibited medication or procedure AR_ETD_03: censoring at the last valid DAA visit at or prior to the PD visit AR_EST_03: censoring at the last valid DAA visit at or prior to the PD visit AR_MD_03: censoring at the last valid DAA prior to the missed visit Remark: The primary analysis of the q12w proportion as derived from DAA and described in section 2.6.2 applies censoring in case the underlying DAA is considered to be confounded by reasons other than lack of efficacy and/or safety. Based on the underlying time-to-'first-q8w-need' analysis, all information up to and including the censoring time-point contribute to the evaluation of the q12w status. **Censoring: subjects are considered to no longer be under risk for a q8w-need identification at later visits.** Censoring at baseline if above PD/AR occurred prior to Week 32. Censoring at Week 52 visit if subjects completed Week 52 without above PD/AR (only applies to Week 52 analysis)
PPS	Censoring of BCVA data after switch to alternative DME treatment in the study eye: imputation using the last observation	Similar to FAS

	collected prior to the start of alternative DME treatment (see section 2.5.3) M_COMD01_Prohibited medication or procedure: censor at the last observation collected prior to the start of the prohibited medication or procedure, imputation using the last observation collected prior to the start of prohibited medication or procedure	

6 References

Bretz F, Maurer W, Brannath W, et al (2009) A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine*; 28(4): 586-604.

Bretz, F., Maurer, W. and Maca J. (2014) Graphical Approaches to Multiple Testing. To appear as Chapter 14 in: *Clinical Trial Biostatistics and Biopharmaceutical Applications* (ed: Walter Young and Ding-Geng (Din) Chen), Taylor and Francis, Boca Raton