Clinical Development

RTH258/Brolucizumab

CRTH258B2305 / NCT03917472

A 12-Month, 2-Arm, Randomized, Double-Masked, Multicenter Phase III Study Assessing the Efficacy and Safety of Brolucizumab every 4 weeks versus Aflibercept every 4 weeks in Adult Patients with Visual Impairment due to Diabetic Macular Edema (KINGFISHER)

Statistical Analysis Plan (SAP)

Author: [redacted], Statistician

Document type: SAP Documentation

Document status: Amendment 1.0 Final

Release date: 28-Jun-2021

Number of pages: 42

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
28-November-2019	Prior to DB lock	Creation of first version	N/A	NA
12-April-2021	Prior to DB lock	Creation of amendment 1	Update definitions of end of treatment and unscheduled visit to make them clearer	Section 2.1.1
			Update analysis restriction and estimand related statement	Section 2.2
			Add subgroup of interest baseline BCVA categories ≤65, >65 letters; Change subgroup of interest baseline status of DRSS score to 12-point scale	Section 2.2.1
			Add analyses to cover potential impact of the COVID-19 pandemic	Sections 2.3.1, 2.5.4.2, 2.5.4.3, 2.6.2, 2.6.2.1, 2.7.1
			Add the summary tables of medical history and prior/concomitant 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, 2.5.4.1, 2.5.4.2
			Change baseline BCVA category factor to ≤ 65, > 65 letters In the ANOVA model	Sections 2.5.2, 2.6.2.1
			Add loss in BCVA in secondary efficacy endpoints	Section 2.6.1
			Update testing strategy to add confirmatory testing on efficacy endpoints	Section 2.6.2.3
			Add handling of censoring for safety analyses	Section 2.7

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Add details on the determination of AESIs	Section 2.7.1.1
			Add section regarding BCVA loss and additional imaging parameters for subjects with AESI in the study eye	Sections 2.7.4.2, 2.7.4.3
			Remove patient profile plot, add NAb analysis	Section 2.9
			Update Table 5-1, Table 5-3	Sections 5.3, 5.5
			Correct minor issues in the first version	

Table of contents

List of abbreviations

ADA	Anti-Drug Antibody
AE	Adverse Event
AESI	Adverse Event of Special Interest
ANOVA	Analysis of variance
AR	Analysis restrictions
AR	First-order Autoregressive
BCVA	Best Corrected Visual Acuity
CI	Confidence interval
CF	Color Fundus
COVID-19	Coronavirus Disease 2019
CS	Compound Symmetry
CSFT	Central subfield thickness (average thickness of circular 1mm area centered around fovea measured from RPE to ILM, inclusively)
CSR	Clinical Study Report
CRC	Central Reading Center
DBL	Database Lock
DME	Diabetic Macular Edema
[redacted]	[redacted]
DRSS	Diabetic Retinopathy Severity Scale
EDTRS	Early Treatment Diabetic Retinopathy Study
EOS	End of Study
EOT	End of Treatment
FA	Fluorescein angiography
FAS	Full Analysis Set
HbA1c	Glycated hemoglobin
hCG	human chorionic gonadotropin
IOP	Intraocular Pressure
IRF	Intraretinal Fluid
IVT	Intravitreal Treatment
KM	Kaplan Meier
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligrams
MMRM	Mixed Model for Repeated Measures
NAb	Neutralizing Antibody
NIM	Non-Inferiority Margin
OCT	Optical Coherence Tomography
PD	Protocol Deviation
PDS	Programming Datasets Specification
PPS	Per Protocol Set
PT	Preferred Term
RAS	Randomized Set

SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SD-OCT	Spectral Domain Optical Coherence Tomography
SE	Standard Error
SOC	System Organ Class
SRF	Subretinal fluid
TEAE	Treatment-Emergent Adverse Event
TOEP	Toeplitz
ULN	Upper Limit of Normal
VA	Visual Acuity
VC	Variance Components
VEGF	Vascular Endothelial Growth Factor
WHO	World Health Organization

1 Introduction

The purpose of this Statistical Analysis Plan (SAP) is to describe the implementation of statistical analysis planned in the study protocol, and to provide detailed statistical methods that will be used for the Clinical Study Report (CSR) of study CRTH258B2305.

Data will be analyzed according to the data analysis Section 12 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 details will be provided, as applicable, in Appendix 16.1.9 of the CSR.

The SAP will be finalized before the 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:

CRTH258B2305 Clinical Trial Protocol Final version 00 dated 26-Mar-2019

1.1 Study design

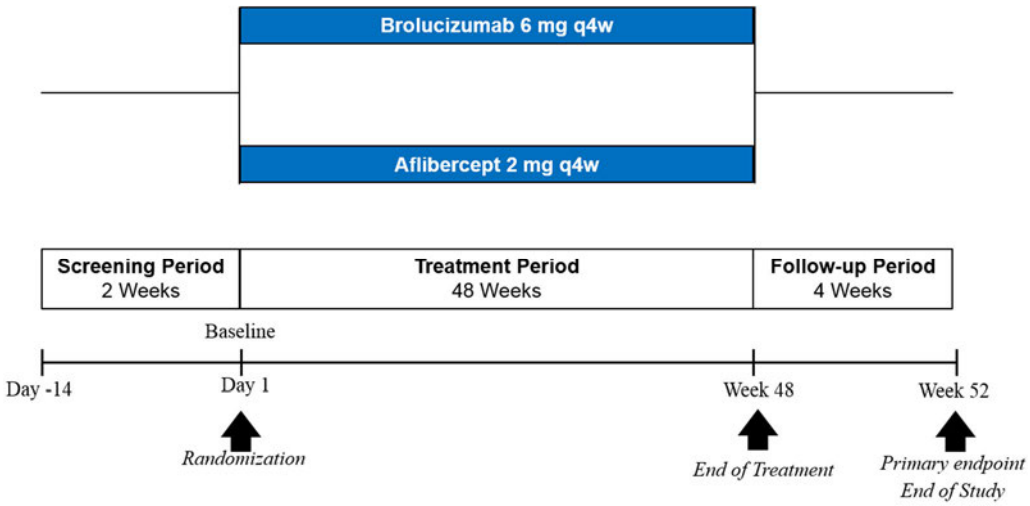
This is a multi-center, randomized, double-masked, parallel group study in subjects with visual impairment due to DME to demonstrate the safety and efficacy of brolucizumab 6 mg q4w against the active control aflibercept 2 mg q4w. Subjects who meet all the inclusion and none of the exclusion criteria will be randomized in a 2:1 ratio to one of two treatment arms:

- Brolucizumab 6 mg: treatment every 4 weeks up to and including Visit 13 (Week 48)
- Aflibercept 2 mg: treatment every 4 weeks up to and including Visit 13 (Week 48)

Approximately 619 adult patients will be screened (20% screening failure rate expected) so that approximately 495 patients will be randomized in a 2:1 ratio (330 in brolucizumab arm, 165 in aflibercept arm, 9% dropout rate expected) in approximately 115 centers worldwide. The study duration is 52 weeks, including follow-up.

All participants will have study visits every 4 weeks through Week 52. The primary analysis will be performed at the End of Study (EOS) visit, Visit 14 (Week 52) (Figure 1-1).

Figure 1-1 Study design



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 non-inferior to aflibercept with respect to the change in visual acuity from baseline up to Week 52	Change from baseline in best-corrected visual acuity (BCVA) at Week 52
Secondary	
To assess the effect of brolucizumab compared with aflibercept with respect to anatomical outcomes	• Change from baseline in CSFT at each post-baseline visit • Proportion of study eyes with fluid-free macula at each post-baseline visit • Proportion of study eyes with absence of DME (CSFT < 280 µm) at each post-baseline visit • Time to first fluid-free macula • Time to first absence of DME (CSFT < 280 µm)
To assess the effect of brolucizumab compared with aflibercept with respect to visual acuity	• Change from baseline in BCVA at each postbaseline visit • Proportion of study eyes with gain in BCVA of 5/10/15 letters or more at each post-baseline visit compared to baseline
To assess the effect of brolucizumab relative to aflibercept on the status of Diabetic Retinopathy	Change from baseline in ETDRS Diabetic Retinopathy Severity Scale (DRSS) score at Week 12, Week 24 and Week 52
To assess the safety and tolerability of brolucizumab compared with aflibercept	Incidence of ocular and non-ocular Adverse Events (AEs)
To assess the immunogenicity of brolucizumab	Anti-drug antibody (ADA) measurement

[redacted]

2 Statistical methods

2.1 Data analysis general information

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 percent. Where appropriate, 2-sided 95% confidence intervals (CIs) for point estimates of the mean or proportion will be provided unless otherwise specified. Point estimates, 95% CIs of treatment differences will be provided as appropriate unless otherwise specified.

2.1.1 General definitions

Study treatment

Study treatment, or study drug, refers to both Brolucizumab 6mg and Aflibercept 2mg 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 since 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 mapping

The end of study 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 end of study 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 maximum change from baseline. Unscheduled visits will not be included in analyses with mixed model for repeated measures (MMRM).

Given unscheduled visits are not active treatment visits, IOP measurements at unscheduled visits are not considered as pre-injection IOP measurements, hence will not be used to identify subjects with pre-injection IOP >30 mmHg.

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** (RAS) will consist of all randomized subjects. Subjects are considered randomized when they had 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. Following the intent-to-treat principle, subjects will be analyzed according to the treatment assigned to at randomization.

Supportive analyses of the primary endpoint will include analysis 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 deviation from inclusion/exclusion criteria. Confounded data or discontinuation from treatment due to lack of efficacy and/or safety do not constitute a reason for exclusion from the PPS.

Before the database lock the relevant important protocol deviations will be identified at the subject level in the database. After the database lock, analysis restrictions will be derived in the analysis database. Censoring applied in relation to the specific PDs / ARs (i.e. non PDs) 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 study drug IVT injection. Subjects in the safety analysis set will be analyzed according to the treatment arm from which they received majority of treatments up to and including Week 48.

Before the database lock, 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.

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 (≤34, >34 letters; ≤65, >65 letters)
- Duration of DME (≤ 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)
- Baseline status of DRSS 12-point scale (≤ 6, ≥ 7)

Subgroup analysis will be performed for the primary efficacy variable only, using the primary analysis approach. More details can be found in Section 2.5.4.

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 are enrolled into the study, treated, complete the

study, discontinue the study (including reasons for discontinuation) and discontinue from study treatment (including reasons for discontinuation).

The number and percent of subjects who discontinue the study and who discontinue treatment will be presented by study visit. The number and percent of subjects treated by site will be presented. A listing of subjects who discontinue 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 sign an informed consent form and who are subsequently found to be ineligible prior to randomization will be considered a screen failure. Screen failure information will not be summarized and only listed.

Number and percent 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 percent 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. The relationship to COVID-19 will be flagged for PDs in the listing.

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. Demographics characteristics will include age, gender, race, 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 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 (≤34, >34 letters; ≤65, >65 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, as defined in Table 2-4).
- Prior anti-VEGF treatment (study eye) (yes, no)

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 related 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 IVT 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
- Treatment exposure by visit: The number and percent of subjects who received IVT injections, missed a treatment and missed visits will be presented by visit

Exposure data will be listed for each treatment arm.

2.4.2 Prior medication and concomitant therapies

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Concomitant therapies are defined as medications and procedures 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) for each treatment arm. Ocular medications will be listed for the study eye and the fellow eye separately.

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.

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

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 objective

2.5.1 Primary endpoint

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

The primary analysis of the primary endpoint will be based on the FAS with last observation carried forward (LOCF) imputation of missing 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

2.5.2 Statistical hypothesis, model, and method of analysis

The objective related to the primary endpoint 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:

B = Brolucizumab 6 mg

A = Aflibercept 2 mg

Consider the following non-inferiority hypotheses related to a non-inferiority margin of 4 letters:

$$\mathbf{H_0}: \mu_B - \mu_A \leq -4 \text{ letters} \quad \text{vs.} \quad \mathbf{H_A}: \mu_B - \mu_A > -4 \text{ letters}$$

where μ_B and μ_A are the corresponding unknown true mean changes from baseline in BCVA at Week 52 in the brolucizumab and aflibercept arms, respectively.

The primary estimand associated with the above hypothesis is defined as the between-treatment difference in change from baseline in BCVA at Week 52, excluding the effect of relevant DME

prohibited medication(s) applied to the study eye (eg, alternative anti-VEGF treatments for DME, as further detailed in protocol Section 6.2.2). The analysis set for the primary estimand will be the FAS.

Based on the FAS, the above hypothesis 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) as factors. Additional factors may also be included as appropriate, as well as interactions between treatment and factors of interest.
Two-sided 95% confidence interval (CI) for the least square means difference (brolucizumab - aflibercept) will be presented. 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. The same approach for non-inferiority assessment in change from baseline in BCVA at Week 52 will be applied to any supplementary estimand.

The primary estimand and other supplementary estimands are described in Table 2-1 below, together with their key attributes.

Table 2-1 Primary and supplementary estimands

Estimand	Estimand definition	Analysis set	Use of data after discontinuation of study treatment due to[1]: use of alternative DME treatment	Use of data after discontinuation of study treatment due to[1]: any other reason	Statistical methods (Including missing data strategy)
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	Not included; treated as missing	Included	Analysis of variance (ANOVA) model assessed at a two-sided significance level of 0.05, and including terms for treatment, baseline BCVA (≤ 65, > 65 letters) and age category (< 65, ≥ 65), and using last observation carried forward (LOCF) imputation/replacement for missing/censored data.
Supplementary Estimand A	Difference in change from baseline in BCVA at Week 52 for subjects adhering to the protocol as per the PPS definition	PPS	Not included; treated as missing	Not Included; treated as missing at subject level[2]	ANOVA model as per the primary estimand. LOCF imputation/replacement for missing/censored data
Supplementary estimand B	Difference in change from baseline in BCVA at Week 52 including the effect of switching to alternative DME medication in the study eye	FAS	Included	Included	ANOVA model as per the primary estimand. LOCF imputation for missing data. No censoring of data after start of alternative DME treatment for the study eye

1 Note that, for all estimands as applicable, all data captured until the start of alternative DME treatment will be included in the analysis;
2 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

As stated in the definition of the primary estimand, 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 treatment but continue in the study, data collected after start of 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 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 outcome 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 discontinuation within the first 6 months of treatment, the LOCF method will likely result in a conservative estimate of the BCVA measure within each arm, 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 analysis 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) with observed data (including censoring for alternative DME treatment in the study eye). The MMRM will include treatment, visit, baseline BCVA category, age category and treatment by visit interactions 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 at Week 52 will be estimated using the LSM and 95% CI.
  - If an MMRM model with 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) and Toeplitz (TOEP), 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 this supplementary estimand will apply the same LOCF/ANOVA method as for the primary estimand.

Supplementary estimand on the FAS:

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. However, data after start of alternative DME treatment in the study eye will be included in the analysis (see Table 2-1).

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 endpoint will use observed data and primary analysis set (FAS), with and without censoring data after use of alternative treatment for DME in the study eye.

Subgroup analyses will be conducted to assess the consistency of treatment effect across various subgroups 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 using LOCF):

- Subgroup analyses will be conducted using the same model and analysis strategies described for the primary analysis 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.
- Subgroup analysis results 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.

Impacted subjects to COVID-19 pandemic are defined as subjects who:

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

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

- did not miss any injection due to COVID-19
- and did not discontinue study treatment due to COVID-19
- and were not reported COVID-19 infection (including suspected as per PTs in PDS).

Subgroup analyses will be conducted using the same model and analysis strategies described for the primary endpoint in the impacted and non-impacted subgroups. In addition,

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

2.6 Analysis of secondary efficacy endpoints

2.6.1 Secondary efficacy endpoints

Secondary efficacy endpoints related to BCVA, anatomy or status of diabetic retinopathy are listed below.

[redacted]

Secondary efficacy endpoints based on BCVA:

- Change from baseline in BCVA at each post-baseline visit
- Proportion of study eyes with gain in BCVA of 5/10/15 letters or more at each post-baseline visit compared to baseline.

  Note: subjects with BCVA value of 84 letters or more at a post-baseline visit will be considered as responders for this endpoint. This is to account for a ceiling effect, for example, for the "≥ 15-letter gain" endpoint, for those subjects with BCVA values at baseline ≥ 70 letters.
- Proportion of study eyes with a loss in BCVA of ≥5, ≥10 and ≥15 ETDRS letters from baseline to each post-baseline visit

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 post-baseline visit
- Proportion of study eyes with fluid-free macula at each post-baseline visit:
  - Proportion of study eyes with absence of subretinal fluid (SRF)
  - Proportion of study eyes with absence of Intraretinal fluid (IRF)
  - Proportion of study eyes with simultaneous absence of SRF and IRF at each post baseline visit

  Note: proportion of fluid-free macula will be based on the full population of FAS.
- Proportion of subject with status regarding normal CSFT thickness (<280 microns) at each post-baseline visit
- Time to first fluid-free macula:
  - Time to first absence of SRF
  - Time to first absence of IRF
  - Time to first absence of SRF and IRF
- Time to first absence of DME (CSFT < 280 microns)

  Note: time to first fluid-free macula should be based on the subset of FAS population with fluid present at baseline.

Secondary efficacy endpoints related to the status of Diabetic Retinopathy (see Section 2.6.2.1**)**:

- Change from baseline in ETDRS Diabetic Retinopathy Severity Scale (DRSS) score at Week 12, Week 24, and Week 52: subject status regarding a ≥2- and ≥3-step improvement or worsening from baseline in the ETDRS DRSS score at each assessment visit

2.6.2 Statistical hypothesis, model, and method of analysis

2.6.2.1 General analysis specifications for secondary efficacy endpoints

No hypothesis will be tested for the secondary efficacy endpoints listed in the above Section 2.6.1.

These endpoints 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 2.6.3.

The impact of COVID-19 pandemic on CSFT will be assessed with the subgroup analyses of 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 including terms for treatment, baseline value categories (BCVA: ≤ 65, > 65 letters; CSFT: <450, 450-650, ≥650 µm) and age category (< 65, ≥ 65). 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 CSFT, baseline CSFT categories will be used instead of baseline BCVA as a class variable.

The line plot on LSMean (±SE) by visit will also be provided for each treatment arm.

Categorical variables:

For binary endpoints, frequency tables (count and proportion) will be provided by time point. In addition, proportions and treatment differences in proportions along with 95% CIs will be presented for each visit 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:

The time to event variables will be analyzed using KM analysis. KM estimates on percent of subjects achieving event together with 95% CI will be presented by visit. The median time (95% CI) to event will also be constructed by treatment arm. KM curves presenting the cumulative probability of subjects achieving event from baseline will be provided by treatment arm. Time to event variables will be analyzed by Cox proportional hazard model, adjusted for treatment, age category and the corresponding baseline value category. The estimated hazard ratio with 95% Wald confidence interval will be obtained.

2.6.2.2 ETDRS DRSS Score

Definition of Endpoints

The following categories of change from baseline in 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

Those endpoints will be derived from the ETDRS-DRSS score assessed by the central reading center based on color fundus photography (CP) images in the study eye at screening, Weeks 12, 24, 52 and exit/premature discontinuation visit. The DRSS in analysis will be from either standard field or the masked DRSS [redacted].

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
60	Full/partial panretinal photocoagulation or local photocoagulation for new vessel
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. Study eyes treated with panretinal photocoagulation or local photocoagulation prior to baseline or during the study period will be coded with score 60 or greater. The improvement or worsening of DRSS in those eyes during study will not truly reflect the treatment effect, so they will be excluded from the DRSS analysis.

All DRSS values except 60 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	DR questionable, microaneurysms only	20

12-point scale	Definition	Original DRSS
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

	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, except for study eyes which were treated with panretinal photocoagulation or local photocoagulation prior to baseline or during the study period.

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 each treatment arm, the difference in proportions between treatment arms and the 95% CI for the difference will be presented for each visit 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.

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 impacted/non-impacted subjects.

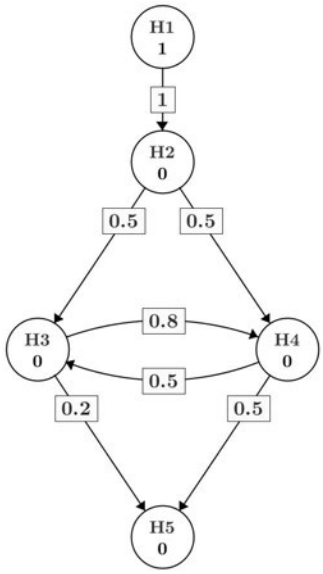
2.6.2.3 Confirmatory testing related to additional secondary efficacy endpoints

Confirmatory hypothesis testing for additional secondary endpoints will be performed in case the proof of non-inferiority related to BCVA is successful for the hypothesis of primary endpoint specified in section 2.5.2 (corresponding to H1 in Figure 2-1).

The additional hypotheses are linked to the endpoints below:

- H2. DRSS ≥ 2-step change from baseline at Week 52 (non-inferior for proportion of subjects in brolucizumab 6 mg vs aflibercept 2 mg with non-inferior margin 10%)
- H3. Absence of SRF and IRF in the study eye at Week 52 (superior for proportion of subjects in brolucizumab 6 mg vs aflibercept 2 mg)
- H4. Change from baseline in CSFT at Week 52 in the study eye (superior in reduction of the CSFT change from baseline in brolucizumab 6 mg vs aflibercept 2 mg)
- H5. Change from baseline in BCVA at Week 52 in the study eye (superior in reduction of the BCVA change from baseline in brolucizumab 6 mg vs aflibercept 2 mg)

Figure 2-1 Multiple testing strategy



- Hypotheses H_1,..., H_5 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 primary hypothesis will be tested.

If the primary hypothesis is rejected at a one-sided significance level of 0.025, the entire alpha will be distributed to the null hypothesis related to the non-inferiority testing of H_2.

If H_2 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_3 (50% of 0.025 = 0.0125), and H_4 (50% of 0.025 = 0.0125).

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 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.

2.6.3 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 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 fluids, 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. If unplanned visit is reported as “Cannot Grade”, no imputation will be applied.

2.7 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 Set.

Except for imputation of partial dates for AEs, no imputations were 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).

2.7.1 Adverse events (AEs)

A treatment-emergent adverse event (TEAE) is defined as any adverse event that develops after initiation of 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) and preferred term (PT). 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 TEAE will be summarized 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		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		Y
[+]≥2 % (or other cutting point as appropriate) in any treatment group for a given PT. # including separate summary tables for impacted/non-impacted subjects to COVID-19.			

In all summary tables listed above, unless otherwise specified, data collected starting from alternative DME treatment after discontinuation of study treatment 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.

2.7.1.1 Adverse events of special interest / grouping of AEs

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

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'.

2.7.1.2 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 the study eye and for the fellow eye will be considered separately.

If for a 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 AE's 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.

2.7.2 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.

2.7.3 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 abnormality 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.

2.7.4 Other safety data

2.7.4.1 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 planned visit will be summarized.

Post-injection IOP are to be assessed within 60 minutes after injection and if ≥25 mmHg, 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 in 3 consecutive scheduled visits will be presented. A visit with missing pre-injection IOP is considered to meet the > 21 mmHg criterion if the preceding 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 >21mmHg 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 ≥10mmHg from pre-injection IOP and a listing of subjects with any IOP > 30mmHg 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.

2.7.4.2 Loss in BCVA

The number and percentage of subjects with a loss in BCVA ≥15, ≥30 letters (study eye) from baseline 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.

2.7.4.3 Vital signs

A summary table with counts and percentage of subjects satisfying the criteria given in Table 5-2 of the Section 5.3 to 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 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.

2.7.4.4 Imaging parameters

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

2.8 Pharmacokinetic endpoints

Not Applicable.

2.9 Anti-drug antibodies

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

ADA status is defined using the following criteria:

- ADA negative:
  - ADA negative at all time points (pre-dose and post-dose)
  - ADA negative at pre-dose and no titer values above 40 at all other time points
  - ADA titer of 40 at pre-dose but negative at all other time points
- ADA positive without boost:
  - ADA positive at pre-dose, post-dose titer values do not increase from pre-dose by more than 3-fold (1 dilution) at any time point
- Induced:
  - ADA negative at pre-dose, post-dose titer value of 120 or more increase
- Boosted:
  - ADA positive at pre-dose, post-dose titer values increase from pre-dose by more than 3-fold (1 dilution) at any time point

The number and percent of subjects according to their integrated ADA status (ADA negative, ADA positive without boost, induced, boosted) will be presented for the brolucizumab treatment arm.

In addition, tabulations by treatment arm will be presented for ADA titer pattern. Subject listings of all ADA titer values will be presented for all subjects.

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.

The number and percentage of subjects with positive neutralizing antibody (NAb) will be summarized by ADA assessment visit for subjects with ADA titer values (ADA positive).

Subject listings of all ADA titer values and NAb status will be presented for all subjects in the brolucizumab arms. Samples collected at unscheduled visits will not be part of the analysis.

In addition to the planned analyses, figures showing the BCVA change from baseline up to Week 52 by ADA status (pre-existing and integrated status up to Week 52) and NAb status (pre-existing and integrated status up to Week 52) will be added. Summary tables of the

incidence of AESIs by ADA status (pre-existing and integrated status up to Week 52) will also be added.

2.11 Interim analysis

There will be no interim analysis in this study.

3 Sample size calculation

The fixed-margin method is used to derive an appropriate non-inferiority margin (NIM) for this study (FDA 2016). Results from a meta-analysis of the phase III studies VIVID and VISTA comparing monthly aflibercept 2 mg vs laser in subjects with DME suggest a pooled mean treatment effect of 10.8 letters (95% CI, [8.0, 13.6]) favoring aflibercept in the change from baseline in BCVA at Week 52. The smallest effect of aflibercept vs laser is defined as the lower limit of the 95% CI for the pooled mean BCVA, ie, an effect of 8.0 letters. The NIM is then set as the value which preserves 50% of this effect, ie, the NIM is set to 4 letters. This noninferiority margin provides assurance that any proof of non-inferiority only occurs if the effect of brolucizumab is superior to that of laser, and the observed treatment difference to aflibercept is of no clinical relevance.

Subjects will be randomized to the brolucizumab and aflibercept arms in a ratio of 2:1. A total sample size of 357 subjects (238 on the brolucizumab arm vs 119 on the aflibercept arm) will allow assessment of non-inferiority (using a NIM of 4 letters) of brolucizumab 6 mg versus aflibercept 2 mg with respect to the change from baseline in BCVA at Week 52. Assuming equal means and a common standard deviation of 11 letters, for a NIM of 4 letters and a two-sided alpha level of 0.05, there is 90% power to reject the null hypothesis that brolucizumab is

inferior to aflibercept. To ensure that at least 300 subjects are treated with brolucizumab 6 mg on a fixed q4w regimen, the total sample size will be increased to 450 subjects (300 on the brolucizumab arm vs 150 on the aflibercept arm). This results in a statistical power for assessing non-inferiority of 95%.

Given the planned unequal sample size of 300 vs 150, there is a higher probability that infrequent AEs will occur in the brolucizumab arm than in the aflibercept arm in this study. For example, for an AE with a probability of occurrence of 0.01, the probabilities of observing at least one event are 95% and 78% for the brolucizumab and aflibercept arms, respectively, using a Binomial distribution.

To account for a dropout rate of 9%, a total of approximately 495 subjects (330 on the brolucizumab arm vs 165 on the aflibercept arm) will need to be randomized.

4 Change to protocol specified analyses

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

Confirmatory hypothesis testing in relation to additional secondary endpoints is introduced in Section 2.6.2.3. Some changes compared to the protocol specified analyses are considered in the current statistical analysis plan before finalization in order to implement the Novartis internal process on SAP simplification (LEAN):

Protocol section	Protocol wording	Change in the SAP
12.3	Descriptive statistics for exposure to study treatment will be provided for the safety set, FAS and PPS if these are different	Descriptive statistics for exposure to study treatment will be provided for the Safety set.
12.4.2	Based on the FAS, the above hypotheses will be tested via an analysis of variance (ANOVA) model. The model will include treatment, baseline BCVA (≤ 34, > 34 letters) and age category (< 65, ≥ 65) as factors. Additional factors may also be included as appropriate, as well as interactions between treatment and factors of interest.	In the ANOVA model, baseline BCVA category factor is changed to ≤ 65, > 65 letters.
12.5.2	**Adverse event of special interest (ESI)** AESI is one of scientific and medical interest to the Sponsor and includes, but is not limited to, the following: • Endophthalmitis • Uveitis: all cases of anterior, posterior, or panuveitis • ⩾ 30 letter decrease in BCVA compared with baseline visual acuity • New retinal tear or detachment	**AESI** AESIs and other safety topics of interest will be identified via the RTH258 electronic case retrieval strategy (eCRS).

	Vital signs All vital signs data will be listed by treatment group, subject, and visit and, if ranges are available, abnormalities will be flagged with thresholds representing clinical relevant abnormality. Summary statistics of absolute and change from baseline data will be provided by treatment and visit. Shift tables using the low/normal/high classification will be used to compare baseline to the worst on-treatment value. **Clinical laboratory evaluations** All laboratory data will be listed by treatment group, subject, and visit, and abnormalities will be flagged (using extended normal ranges as provided by the central laboratory) with thresholds representing clinical relevant abnormality. Summary statistics of absolute and change from baseline data will be provided by treatment and visit. Shift tables using the low/normal/high classification will be used to compare baseline to the worst on-treatment value.	**Vital signs** A summary table with counts and percentage of subjects satisfying the criteria given in Table 5-2 of the Section 5.3 to 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 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. **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 abnormality 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.

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. 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)
    b. else if DIAG month is not missing, the imputed DIAG date is set to the mid-month point (15MONYYYY)
2. 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

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

5.2 AEs coding/severity

AEs are coded using the MedDRA terminology.

AEs severity are 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-numeric
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

In the sample SAS code, the terms in the brackets are for demonstration purpose and to be adjusted according to the real data set and variable names.

5.4.1 Primary analysis

The primary endpoint (change from baseline in BCVA at Week 52) will be analyzed using ANOVA and MMRM models.

Analysis of Variance (ANOVA)
The following ANOVA model will be used for the primary endpoint:
<change from Baseline in BCVA at 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 efficacy variable:
*<change from Baseline in BCVA at 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 ANOVA model for continuous variables

The continuous efficacy variable (such as CSFT change from baseline at Week 52) will be analyzed using an ANOVA model adjusted for treatment, age category, and the corresponding baseline value category.
The SAS Proc MIXED will be used to perform the ANOVA 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 corresponding covariates if necessary, on 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.
- For superiority tests, p-values (if applicable) will be obtained from odds ratio test in logistic model. For non-inferiority tests, p-values (if applicable) will be obtained by applying the bootstrap method.

5.4.2.3 Cox and KM estimate for time to event variables

Time to event parameters will be analyzed by Cox proportional hazard model, adjusted for treatment, age category and the corresponding baseline value category. The estimated hazard ratio with 95% Wald confidence interval will be obtained. SAS Proc Phreg will be used to perform the Cox model.

The median time and 95% CI will be derived from KM method, using SAS Proc Lifetest.

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.

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

Deviation ID	Description of Deviation	Exclusion in Analyses
M_INCL01_ICF not obtained	Written informed consent not obtained	Exclude from all analysis
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 12% at screening	Exclude from PP analysis
P_INCL04a_No visual impairment study eye	Study Eye: no visual impairment due to DME as per BCVA criteria	Exclude from PP analysis
M_INCL04b_No visual impairment study eye	Study Eye: no visual impairment due to DME as per CSFT criteria	Exclude from PP analysis
M_EXCL08_ condition with impact study eye	Study Eye: Confounding ocular concomitant conditions or ocular disorders with impact on efficacy and/or safety	Exclude from PP analysis
M_EXCL10_Confounding treatment study eye	Study Eye: Confounding concomitant medications or procedures	Exclude from PP analysis
M_EXCL11_Confounding systemic condition/trt with impact	Systemic: confounding systemic conditions or systemic treatments with impact on efficacy and/or safety	Exclude from PP analysis
M_TRT01_Wrong IP administered	Wrong IP administered during the study	Exclude from PP analysis
M_TRT02_Under-treatment	Under-treatment: Missed 2 doses (not due to any safety event)	Exclude from PP analysis
M_TRT03_Over treatment	Over treatment: 3 doses < 41 days	Exclude from PP analysis
M_OTH01_Other PD with impact	Any other protocol deviation with impact on efficacy and/or safety	Exclude from PP analysis

Deviation ID	Description of Deviation	Exclusion in Analyses
M_OTH03_T EXC IP	Administered IP that underwent a temperature excursion	Exclude from PP analysis
P_WITH01_Treatment but consent withdrawn	Subject withdrew consent but continue in the study and 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 visits / missed treatments).

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 analysis
AR_EST_02	Early study termination due to safety	2	Include in all analysis
AR_EST_03	Early study termination due to reasons other than lack of efficacy/safety before or at week 48	0	Exclude from PP analysis
AR_ETD_01	Early study treatment termination due to lack of efficacy	1	Include in all analysis
AR_ETD_02	Early study treatment termination due to safety	2	Include in all analysis
AR_ETD_03	Early treatment termination due to reasons other than lack of efficacy/safety before or at week 48	0	Exclude from PP analysis

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.

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
RAS	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_INCL04a_No visual impairment study eye, M_INCL04b_No visual impairment study eye, M_EXCL08_ condition with impact study eye, M_EXCL10_Confounding treatment study eye, M_EXCL11_systemic condition/trt with impact M_TRT01_Wrong IP administered, M_TRT02_Under-treatment, M_TRT03_Over treatment, M_OTH01_other PD With Impact, P_WITH01_Treatment but consent withdrawn, M_OTH03_T EXC IP	Not in the FAS, AR_EST_03, AR_ETD_03

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 endpoint are described in Section 5.5.

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

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

Analysis Set	Censoring concept for BCVA
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.
PPS	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)

Analysis Set	Censoring concept for BCVA
	M_COMD01_Prohibitedtrt with impact: 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 Reference

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.