Clinical Development and Regulatory Affairs

Biostatistics and Data Management

REGENERON
Biostatistics and Data Management

STATISTICAL ANALYSIS PLAN

Title:	RANDOMIZED, CONTROLLED, MULTI-CENTER STUDY TO ASSESS THE EFFICACY, SAFETY, AND TOLERABILITY OF INTRAVITREAL AFLIBERCEPT COMPARED TO LASER PHOTOCOAGULATION IN PATIENTS WITH RETINOPATHY OF PREMATURITY
Protocol Number:	VGFTe-ROP-1920 Amendment 1 29Jul2019
Investigational product:	Intravitreal Aflibercept Injection
Sponsor:	Regeneron Pharmaceuticals, Inc.
Study Biostatistician:	[redacted]
Clinical Trial Manager:	[redacted]
Study Medical Director:	[redacted]
Date:	**05MAY2021**
Version:	**4.0**

The approval signatures below indicate that these individuals have reviewed the Statistical Analysis Plan (SAP) and agreed on the planned analysis defined in this document for reporting.

See appended electronic signature page

I Approve this Document

[redacted]

See appended electronic signature page

I Approve this Document

[redacted]

See appended electronic signature page

I Approve this Document

[redacted]

See appended electronic signature page

I Approve this Document

[redacted]

See appended electronic signature page

I Approve this Document

[redacted]

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AE	Adverse Event
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutically Chemical
BUN	Blood Urea Nitrogen
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
DMC	Data Monitoring Committee
DWFI	Digital Wide-Field Images
EOS	End of Study
FA	Fluorescein Angiography
FAS	Full Analysis Set
FP	Fundus Photography
ICF	Informed Consent Form
ICH	International Conference of Harmonization
IOP	Intraocular Pressure
IVT	Intravitreal
LDH	Lactate Dehydrogenase
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
(MedDRA) HLT	High Level Term
(MedDRA) LLT	Low Level Term
(MedDRA) PT	Preferred Term
(MedDRA) SOC	System Organ Class
PPS	Per Protocol Set
RBC	Red Blood Cell
RC	Reading Center
ROP	Retinopathy of Prematurity
SAE	Serious Adverse Event
SAF	Safety Analysis set
SAP	Statistical Analysis Plan

TEAE	Treatment-emergent Adverse Event
UPCR	Urine Protein Creatinine Ratio
VEGF	Vascular Endothelial Growth Factor
VEGFR1	Vascular Endothelial Growth Factor Receptor 1
VEGFR2	Vascular Endothelial Growth Factor Receptor 2
VEGF-A	Vascular Endothelial Growth Factor-A
WBC	White Blood Cell
WHO	World Health Organization

DOCUMENT VERSION HISTORY

Version	Date	Version History
1.0	02OCT2019	Initial Version
2.0	21FEB2020	Purpose: Update to incorporate the following revisions in response to feedback from the Food and Drug Administration (FDA): Section 3.4: Add the population of All Randomized Set Section 5.5.1.1: Define that patients who discontinue due to an adverse event between Week 40 and Week 52 will be considered non-responders. Section 5.5.1.2: Add a sensitivity analysis for the primary variable to be performed based on all randomized patients.
3.0	13DEC2020	Purpose: Update to incorporate the following revisions in response to feedback from the Food and Drug Administration (FDA): Section 2.1 Add the power for one-sided 0.025 significance level. Section 5.5.1.1 Add one-sided 97.5% confidence intervals as supportive analysis.
4.0	05MAY2021	Purpose: Update to incorporate the following revisions in response to feedback from FDA: Section 1.2.4. Document the changes from the statistical section in the protocol Section 5.1. Change the significant level to 2.45% (one-sided test) (reflecting an alpha adjustment of 0.001 (two-sided) for the DMC assessments) Section 5.5.1.1. Change confidence interval from one-sided 95% to two-sided 95.1% and remove supportive analysis of one-sided 97.5% confidence interval. Section 5.5.4. Clarify the testing hierarchy for superiority test of primary endpoint Section 7. Add the alpha adjustment for Data Monitoring Committee (DMC) safety assessments.

1. OVERVIEW

The purpose of the statistical analysis plan (SAP) is to ensure the credibility of the study results by pre-specifying the statistical approaches for the analysis of the study. The statistical evaluation will be done according to the specifications given in the protocol and, if applicable, the corresponding amendments. This SAP covers analyses through end of study visit (52-weeks of chronological age) only.

1.1. Background/Rationale

As stated in the study protocol, Retinopathy of prematurity (ROP) is a proliferative vascular retinopathy caused by an abnormal development of the vascularization of the peripheral retina in premature infants. It mainly affects newborns with a preterm gestational age (≤32 weeks) and very low birth weight (≤1500 g). Approximately 10% to 11% of all newborn infants are born preterm globally. Among premature infants, the incidence of ROP ranges from 20% to 30%, increasing for younger gestational ages. The prevalence of blindness caused by ROP in developed nations ranges between 6% and 18%, while in developing nations the estimate is >20%. Epidemiological studies have also identified increased risks for ROP due to genetic variants and environmental factors, such as oxygen exposure.

Vascular endothelial growth factor (VEGF) is up-regulated under ischemic conditions. As ROP is characterized by incomplete vascularization of the retina in premature infants, it has also been associated with increased levels of VEGF. Vascular endothelial growth factor-A (VEGF-A) is a member of the VEGF family of angiogenic factors that acts through activation of the receptor tyrosine kinases vascular endothelial growth factor receptor 1 (VEGFR1) and vascular endothelial growth factor receptor 2 (VEGFR2) to increase mitosis, chemotaxis, and vascular permeability in endothelial cells. Upregulation of VEGF-A due to ischemia in the avascular retina may induce pathologic neovascularization and subsequently lead to retinal detachment and blindness, as seen in late-stage patients. Aflibercept inhibits these effects of VEGF by acting as a soluble decoy receptor that binds VEGF-A with higher affinity than the natural receptors and thereby inhibits the binding and activation of these cognate VEGF receptors. Aflibercept also inhibits the effects of placental growth factor, which binds only to VEGFR1. Placental growth factor can synergize with VEGF-A, enhancing these processes of pathologic neovascularization, and is also known to promote leukocyte infiltration and vascular inflammation. Based on the its mode of action, aflibercept has a high potential to become an effective treatment option for the treatment of ROP, and several reports of off-label clinical use have shown positive efficacy outcomes without indicating new safety concerns.

This Phase 3, multicenter, randomized, 2-arm open-label study will assess the efficacy, safety, and tolerability of intravitreal (IVT) aflibercept compared to laser in patients (premature infants) diagnosed with ROP.

1.2. Study Objectives

1.2.1. Primary Objectives

The primary objective of the study is to assess the efficacy of aflibercept compared to laser in patients diagnosed with ROP.

1.2.2. Secondary Objectives

The secondary objectives of the study are:

- To assess the need for a second treatment modality
- To assess the recurrence of ROP in the study
- To assess the safety and tolerability of aflibercept

1.2.3. Exploratory Objectives

The exploratory objectives of the study are:

- To further characterize the anatomical effects of treatment with aflibercept in patients with ROP
- To further investigate the study intervention (eg, mode-of-action-related effects and/or safety) relevant to this disease process
- To assess the treatment burden of aflibercept and laser
- To characterize the concentrations of free and bound aflibercept in plasma over time
- To describe the potential immunogenicity of aflibercept

1.2.4. Modifications from the Statistical Section in the Final Protocol

To incorporate the feedback from FDA, the followings are modified from the statistical section in the final protocol.

For the analysis of primary efficacy variable, the confidence interval is updated from one-sided 95% to two-sided 95.1%, An alpha adjustment of 0.001 is made to the overall alpha at the final analysis in order to account for the DMC safety assessments.

1.2.5. Modifications from the Approved Statistical Analysis Plan

Not applicable

2. INVESTIGATION PLAN

2.1. Sample Size and Power Considerations

The RAINBOW [1] study compared Ranibizumab 0.1 mg and 0.2mg to Laser, in the treatment of ROP. The primary endpoint is the absence of both active ROP and unfavorable structural outcomes at 24 Weeks. From the RAINBOW study, the response rate was 66.1% for the laser group and 88.1% for the 0.2 mg ranibizumab group in zone II disease. Furthermore, based on the clinical evidence for the aflibercept investigator-initiated studies (Salman, 2015[2]) (Huang, 2018[3]) (Sukgen, 2018[4]) (Sidorenko, publication in progress), the response rates range all the way up to 100% (with at least 2 studies demonstrating a 100% response rate in terms of favorable anatomic structural outcomes/prevention of unfavorable anatomic structural outcomes) with intravitreal aflibercept doses ranging from 0.4 mg to 1 mg. Given the above data, an

estimated response rate for the primary efficacy variable of 90% for the aflibercept group and 66.1% for the laser group is considered to be a reasonable assumption.

A sample size of 84 patients in the aflibercept group and 28 patients in the laser group (randomized in a 3:1 ratio) will provide 92% power for testing non-inferiority (NI) of aflibercept versus laser at a one-sided 0.05 significance level assuming non-inferiority margin of 5% (See Section 5.1 for statement of the statistical hypotheses and justification of the NI margin). In addition, this sample size (a total of 112 patients randomized in a 3:1 ratio to aflibercept injection or laser, respectively) will provide 86% power for one-sided 0.025 significance level.

The sample size calculation was computed using the binomial non-inferiority two sample test from the commercial software EAST version 6.4.1.

2.2. Study Design and Randomization

This is a phase 3, multicenter, randomized, 2-arm, open-label clinical study to assess the efficacy, safety, and tolerability of intravitreal (IVT) aflibercept versus laser in patients with ROP. The study consists of screening/baseline (1 or 2 visits), a treatment period (including potential retreatment and rescue treatment), and a final visit at week 52 of chronological age (Figure 1).

The study also includes an optional fluorescein angiography (FA) sub-study, which includes optional FA at baseline, and mandatory FA at week 52 of chronological age.

Successful screening requires the presence of treatment-naïve ROP classified by the investigator according to the International Classification for Retinopathy of Prematurity in at least 1 eye with 1 of the following retinal findings:

- Zone I Stage 1 plus, 2 plus, 3 non-plus or 3 plus, or
- Zone II Stage 2 plus or 3 plus, or
- AP-ROP

One or both eyes can be treated according to the investigator's assessment of the study's eligibility criteria. If both eyes are eligible, they will be assigned to the same treatment group. The patient who starts the study with only 1 eligible eye should be kept under observation according to the local ROP screening guidelines or at every study visit, whichever is more frequent. Second eyes that develop ROP requiring treatment during the study should receive treatment according to the randomization assignment of the first eye. If the second eye is treated within 8 weeks of the initial treatment of the first eye, the second eye will be included in the efficacy analysis.

For the primary efficacy endpoint, the patient will be the experimental unit. If both eyes are eligible and treated within 8 weeks, the success of the primary efficacy endpoint will be determined by both eyes. If both eyes are eligible but the second eye is treated outside 8 weeks of the initial treatment of the first eye, or only one eye is eligible, then the success of the primary endpoint will be determined by one eye.

Color fundus photography (FP) with digital wide-field images (DWFI) of the retina will be taken before any treatment is applied. These images will be submitted to a central reading center (RC) as soon as possible for confirmation of ROP staging, and if deemed eligible, treatment should be

administered. However, due to the urgency of treatment for this medical condition, in situations where the investigator considers that awaiting a response from the RC and thus delaying treatment may be detrimental to the outcome, treatment is allowed to be administered immediately after the images are acquired, before availability of the RC confirmation of ROP staging. In this case, the investigator's assessment of the staging will be used in stratification.

Patients will be randomized 3:1 to treatment with either aflibercept injection or laser, respectively, stratified by ROP classification in Zone I, Zone II, or AP-ROP according to investigator assessment.

Aflibercept group:

Patients randomized to aflibercept will receive a single IVT injection of aflibercept 0.4 mg/0.01 mL per eligible eye at baseline.

Thereafter, if required, up to 2 additional IVT injections of aflibercept 0.4 mg/0.01 mL may be administered in each eye if the following retreatment criteria are met:

- Presence of ROP requiring treatment AND
- The interval since the last aflibercept IVT injection is ≥28 days

Rescue treatment with laser may be performed if 1 of the following conditions is met:

- Worsening of ROP compared to the examination immediately preceding the previous aflibercept injection, during the 27 days following the aflibercept injection
- Presence of ROP requiring treatment after the patient already received a total of 3 aflibercept injections

Laser Group:

Patients randomized to laser will undergo treatment in each eligible eye at baseline. Laser ablation should be as complete as possible as judged by the investigator. In case multiple sessions are necessary within 1 week from baseline, they will be counted as a single treatment. Treatment will be applied to the entire avascular peripheral retina. Treatment should be kept well away from the fovea.

Supplementary laser treatments are allowed during the study. Retreatment with laser is allowed if both of the following criteria are met:

- Presence of ROP requiring treatment
- Fundus examination reveals that laser treatment is incomplete as judged by the investigator

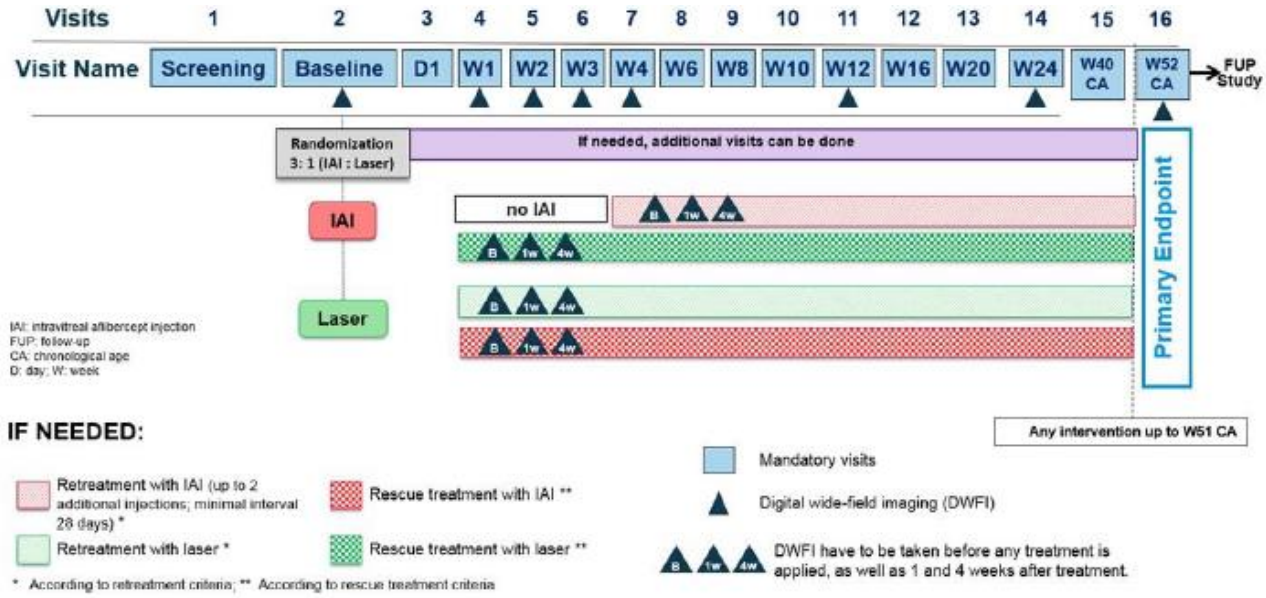
Rescue treatment with aflibercept 0.4 mg/0.01 mL is allowed if the fundus examination reveals that laser treatment is complete as judged by the investigator and if 1 of the following conditions is met:

- Worsening of ROP compared to the most recent pre-laser examination
- Persistence of ROP requiring treatment

Patients who initiate aflibercept rescue treatment may receive additional aflibercept injections according to the aflibercept group treatment regimen.

The study event table is presented in Appendix 10.1.

Figure 1: Study flow diagram



3. ANALYSIS POPULATIONS

In accordance with guidance from the International Conference of Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guideline ICH E9 Statistical Principles for Clinical Trials (ICH,1998 [5]), the following analysis populations will be used for all statistical analysis:

3.1. Full Analysis Set

The full analysis set (FAS) will include all randomized patients who received any study treatment. The analysis on the FAS will be performed according to the treatment assigned at baseline (as randomized).

3.2. Safety Analysis Set

The safety analysis set (SAF) includes all randomized patients who received any study treatment (active or laser); it is based on the treatment actually received (as treated). Treatment compliance/administration and all clinical safety variables will be analyzed using the SAF.

3.3. Per Protocol Set

The Per Protocol Set (PPS) includes all patients in the FAS who had no important protocol deviations. Analysis of the PPS will be performed according to the treatment actually received

(as treated). The PPS will be used for sensitivity analysis of primary efficacy endpoint. The protocol deviation plan will be finalized prior to start of study enrollment.

3.4. All Randomized Set

All Randomized Set will include all randomized patients regardless receiving study treatment or not. The analysis on the all randomized patients will be performed according to the treatment assigned at baseline (as randomized).

3.5. Pharmacokinetic Analysis Set

The Pharmacokinetic (PK) Analysis Set includes all patients who received any study treatment and who had a at least 1 non-missing result following the first dose of study treatment.

3.6. Immunogenicity Analysis Sets

The ADA analysis set (AAS) includes all patients who received any study treatment and had at least one non-missing ADA result from the Aflibercept ADA assay after first dose of the study treatment. Patients will be analyzed according to the treatment actually received.

4. ANALYSIS VARIABLES

4.1. Demographic and Baseline Characteristics

Demographic and baseline assessments to be summarized will include:

- Chronological age at randomization (date of randomization - date of birth + 1)
- Gestational age at birth (in weeks and days)
- Post-menstrual age at randomization (Gestational age at birth + Chronological age at randomization)
- Gender
- Race
- Body weight, head circumference and body length at birth and at baseline
- Vital Signs (Baseline pulse rate, systolic blood pressure, diastolic blood pressure and temperature)
- Baseline Intraocular pressure (IOP)
- ROP classification (Zone I, Zone II, plus status and AP-ROP)
- O2 supplementation at baseline (yes, no)
- History of sepsis (yes, no)
- History of necrotizing enterocolitis (yes, no)
- History of intraventricular hemorrhage (yes, no)

- Number of patients with two/one eye(s) enrolled

4.2. Medical History

Maternal and patient Medical history will be coded according to latest available version of Medical Dictionary for Regulatory Activities (MedDRA).

4.3. Pre-Treatment / Concomitant Medication

Medications taken during the study will be recorded and will be coded to Anatomical Therapeutically Chemical (ATC) codes according to the World Health Organization Drug Dictionary (WHO Drug Dictionary) B3 Global 201903 enhanced version provided by Bayer Health Care.

Medications will be summarized as follows:

- Maternal prior medications taken during pregnancy (i.e. medication taken by the mother with start day prior to date of birth of child but after conception)
- Prior medication taken by the breastfeeding mother (i.e. medication taken by the mother with start date after date of birth but prior to the start of study treatment and end date prior to the start of study treatment)
- Prior medication taken by the patient
- Concomitant medication taken by the breastfeeding mother
- Concomitant medication taken by the patient

The prior, concomitant and new medication will be summarized by ATC class (ATC level 1) and subclass (ATC level 2).

Variables for concomitant medication description and analysis will include Generic name, ATC level codes, Indication, Dose/Dose Unit, Frequency, Route, start/end date and study day, Duration, Ongoing.

4.4. Exposure and Additional Treatment to Study Treatment

Exposure

For each patient, the following variables will be used to examine exposure to study treatment (rescue treatment will not be taken into account here):

- Total number of treatment (aflibercept or laser) by treatment group
- For patients who receive rescue treatment, frequency of treatment will be summarized
- Time required to perform treatment
- Requirement for sedation or general anesthesia
- Injection volume
- Number of burns, wave length, spot sizes, power and duration of laser pulse (for laser treatment)

4.5. Efficacy Variable

4.5.1. Primary Efficacy Variable (s)

The primary efficacy variable is the proportion of patients with the absence of both active ROP and of unfavorable structural outcomes at 52 weeks of chronological age, as determined by the Investigator. For patients with both eyes enrolled in the study, both eyes must meet the endpoint.

Active ROP is defined as ROP requiring treatment (according to the inclusion criteria). Unfavorable structural outcomes are defined as retinal detachment, macular dragging, macular fold, or retrolental opacity.

Furthermore, patients will be considered as non-responders if rescue treatment is given in the study eye(s).

4.5.2. Secondary Efficacy Variable(s)

The secondary efficacy variables are:

- Proportion of patients requiring intervention with a second treatment modality from baseline to week 52 of chronological age
- Proportion of patients with recurrence of ROP through week 52 of chronological age. Recurrence of disease is defined further in Section 5.5.2.

4.5.3. Exploratory Efficacy Variable(s)

The exploratory efficacy variables are:

- Time to recurrence of ROP
- Time to intervention with a second treatment modality for ROP or to development of unfavorable structural outcomes
- Proportion of patients with completion of vascularization of the peripheral retina to within 1 disc diameter of the ora serrata on FP at week 52 of chronological age
- Presence of retinal vascularization and leakage on fluorescein angiography (FA) at week 52 of chronological age in patients participating in the FA substudy
- Time to completion of retinal vascularization
- Proportion of patients with regression of plus disease, regression of pre-retinal vascularized ridge, and progression of retinal vascularization beyond the ridge from baseline to week 52 of chronological age
- Proportion of patients with progression to stage 4 or 5 ROP from baseline to week 52 of chronological age
- Evaluation of visual function at week 52 of chronological age
- Time required to perform aflibercept injection or initial laser treatment
- Proportion of patients with requirement for sedation or general anesthesia to complete laser or aflibercept injection

4.6. Safety Variables

4.6.1. Adverse Events and Serious Adverse Events

An adverse event (AE) is any untoward medical occurrence in a patient administered a study treatment which may or may not have a causal relationship with the study treatment.

Adverse events will be collected at each visit from the time of informed consent signature until the end of the study. If the patient withdraws from the study during the screening, AEs will be collected up until the patient withdraws. If the patient withdraws at any point after receiving the first dose of study treatment, AEs will be collected up until 30 days after the last dose of study treatment or the termination visit, whichever is later.

The focus of AE reporting in the clinical study report will be on **Treatment-Emergent Adverse Events (TEAEs) :** TEAEs are defined as AEs that developed or worsened during the on-treatment period, that is, AEs that are observed or reported after first and not later than 30 days after last administration of study treatment (laser or aflibercept injection) in the study eye(s), laser or aflibercept rescue treatment in the fellow eye.

Other variables for AE description and analysis will include AE Verbatim Term, AE start date and end date/ongoing and corresponding study day, AE Duration, relationship of AE to study treatment, relationship of AE to study conduct, seriousness, intensity, action due to AE, treatment of AE and outcome.

4.6.2. Surgeries

All the surgeries after informed consent are collected on the CRF and are coded by MedDRA.

The following variables will be tabulated by MedDRA preferred term:

- Treatment emergent surgery is defined as surgery performed on or after the start of study treatment
  - Ocular treatment emergent surgery for study eye(s) and fellow eye
  - Non-ocular treatment emergent surgery

4.6.3. Laboratory Safety Variables

Clinical laboratory variables will include the following:

- Blood chemistry panel: Sodium, Total protein, serum, Total bilirubin, Potassium, Creatinine, Total cholesterol, Chloride, Blood urea nitrogen (BUN), Triglycerides, Carbon dioxide, Aspartate aminotransferase (AST), Uric acid, Calcium, Alanine aminotransferase (ALT), Creatine phosphokinase (CPK), Glucose, Alkaline phosphatase, Albumin, Lactate dehydrogenase (LDH)
- Hematology panel: Hemoglobin, Hematocrit, Red blood cells (RBC), White blood cells (WBC), Red Cell Indices, Platelet count, Differential: Neutrophils, Lymphocytes, Monocytes, Basophils, Eosinophils

- Urinalysis: Color, Glucose, RBC, Clarity, Blood, Hyaline and other casts, pH, Bilirubin, Bacteria, Specific gravity, Leukocyte esterase, Epithelial cells, Ketones, Nitrite, Crystals, Protein, WBC, Yeast

4.6.4. Vital Signs

Variables of analysis for vital signs include temperature, pulse, blood pressure, and respiration rate.

4.6.5. Ocular Safety Measures

Variables of analysis for ocular safety measures include:

- Proportion of patient with increased intraocular ocular pressure (IOP)
  - ≥ 10 mmHg increase in IOP measurement from baseline to any pre-dose measurement
  - > 21 mmHg for any pre-dose measurement
  - ≥ 25 mmHg for any pre-dose measurement
  - ≥ 35 mmHg at any time

Post dose IOP measurement should be the last IOP recorded.

4.7. Pharmacokinetic Variables

The PK variable is the concentration of free or bound aflibercept in plasma at each time point.

4.8. Immunogenicity Variables

The immunogenicity variables are ADA status, titer, and time-point/visit. Serum samples for ADA will be collected at the clinic visits specified in Appendix 10.1.

5. STATISTICAL METHODS

All efficacy and safety variables will be summarized descriptively with appropriate statistics: categorical variables by frequency (absolute and relative frequencies) and continuous variables by sample statistics (i.e. mean, standard deviation, minimum, median, quartiles and maximum). Continuous variables will be described by visit and as change from Baseline, if applicable.

5.1. Statistical Hypothesis

This study will examine the following hypothesis for the primary efficacy variable regarding the proportion of patients with both the absence of active ROP and unfavorable structural outcomes at 52 weeks of chronological age. Statistical testing will be conducted to demonstrate the non-inferiority of the aflibercept group to the laser group, with a non-inferiority margin of 5% and a significance level of 2.45% (one-sided test) (reflecting an alpha adjustment of 0.001 (two-sided) for the DMC assessments).

H_0: $p_t \leq p_c$ - 5% versus H_1: $p_t \geq p_c$ - 5%

where p_t is the true proportion of patients with absence of active ROP and unfavorable structural outcomes at week 52 of chronological age for the aflibercept treated group, and p_c is the true proportion of patients with absence of active ROP and unfavorable structural outcomes at week 52 of chronological age for the laser control group.

In this trial, laser is the active control arm. From the RAINBOW study, the success rate for laser was 66.2% (95% CI: 55.0% to 77.4%), while the success rate for a putative placebo (which has little to no efficacy) is assumed to be near 0%. The 95-95 approach described in Rothman [6,7,8] was used to determine the non-inferiority (NI) margin for the study. Using this method, the margin, M1 (also known as statistical margin), is conservatively assumed to be the lower limit of 95% CI for the treatment difference (laser – putative placebo), i.e., 55%. The smaller margin, M2=5% (also known as the clinical margin) was prespecified such that a large fraction of the active control (laser) treatment effect is preserved for treating ROP. The proposed NI margin of 5% is the smaller of M1 and M2, and preserves at least 90.9% of the control treatment effect. Therefore, this NI margin is adequate and justified.

5.2. Subject Disposition

The following categories for patient disposition will be summarized descriptively:

- The total number of screened patients: met the inclusion criteria regarding the target indication and signed the informed consent form (ICF)
- The total number and percentage of randomized patients: received a randomization number
- The number and percentage of patients in each analysis set
- The total number and percentage of patients who discontinued the study with the reasons for discontinuation

The following listings will be provided to assess the patient disposition:

- A listing of patients treated but not randomized and patients randomized but not treated if any
- A listing of patients who received rescue treatment in the study eye(s)
- A listing of patients who were withdrawn from the study, along with reasons for discontinuation
- Listing of important protocol deviations: violation of inclusion/exclusion criteria; post-enrollment deviations which will impact assessment of efficacy endpoints

5.3. Demographics, Baseline Characteristics, Medical History, and Prior/ Concomitant Medication

Demographic data and baseline characteristics variables described in Section 4.1 will be summarized using descriptive statistics for SAF and FAS.

Medical history is evaluated for SAF by a frequency table, showing number of patients with medical history findings by primary system organ class (SOC), high level term (HLT) by MedDRA terms.

Prior/concomitant medication will be summarized by WhoDrug B3 Global 201903ATC codes (ATC 3-digit class and ATC 5-digit subclass) for medication taken during the study. Separate frequency tables will be displayed for patients with prior medications, new medications and concomitant medications by the time periods described in Section 4.3.

5.4. Extent of Study Treatment Exposure

The variables for dose exposure described in Section 4.4 will be summarized in the SAF and FAS populations, using descriptive statistics.

5.5. Analyses of Efficacy Variables

Efficacy analyses of all efficacy variables defined in Section 4.5 will be conducted using the FAS population. PPS population will be used as sensitivity analysis for primary efficacy variable.

5.5.1. Analysis of Primary Efficacy Variable(s)

5.5.1.1. Primary Analysis for Primary Efficacy Variable

The primary efficacy variable analysis will be conducted on the FAS. The primary analysis is a statistical evaluation of non-inferiority of aflibercept vs. laser at week 52 of chronological age, with respect to the primary efficacy variable (i.e., absence of both active ROP and of unfavorable structural outcomes; see Section 4.5.1). The non-inferiority margin is set at 5%. Superiority testing will follow if non-inferiority is established.

If the patient data are not available at end of study (EOS) visit, then data available from the week 40 of chronological age visit will be carried forward (LOCF) to the EOS visit for analysis except the following two cases:

- If the patient discontinues at or before the week 40 visit, no data will be carried forward and the patient will be considered a non-responder.
- If patients discontinue the study due to AEs between week 40 and week 52, then these patients will be considered as non-responders.

For patients with both eyes enrolled in the study, both eyes must meet the endpoint. Patients with only one study eye enrolled will be responders if the respective eye responds.

The statistical analysis will be performed using the Cochran-Mantel-Haenszel method stratified by baseline ROP status. The two-sided 95.1% Mantel-Haenszel confidence intervals (reflecting an alpha adjustment of 0.001 for the DMC assessments) using normal approximation of the difference of response rates between the aflibercept group and the laser group will be calculated. Aflibercept will be considered to be non-inferior to laser if the lower confidence bound of the difference lies above -5%. Furthermore, if the lower confidence bound lies above 0%, superiority of Aflibercept may be declared.

The primary analysis for this endpoint will be based on the investigator assessment.

The components of the primary efficacy variable, i.e., number and percentage of patients with absence of active ROP and/or number and percentage of patients with unfavorable outcomes will also be presented.

5.5.1.2. Sensitivity Analyses for Primary Efficacy Variable

The primary efficacy variable will be analyzed using the PPS population and all randomized population as sensitivity analyses.

Observed case (OC) analysis will be conducted as supplementary analysis. Only observed values will be used for analysis. Patients will be considered as non-responders if rescue treatment is given.

The same analysis for the primary endpoint will be performed based on the central RC data as a sensitivity analysis.

5.5.2. Analysis of Secondary Efficacy Variables

The following secondary efficacy variables will be analyzed using the same method as for the analysis of the primary efficacy variable. Patients will be counted as "event" if at least one eye satisfies the criteria.

- Proportion of patients requiring intervention with a second treatment modality from baseline to week 52 of chronological age
- Proportion of patients with recurrence of ROP through week 52 of chronological age

Recurrence of disease is defined as the reappearance of the disease requiring further treatment (including retreatment or rescue), where both "presence of ROP" and "presence of active ROP requiring treatment" are marked as "Yes", after initial regression. Here, the initial regression is defined as, at particular visit, absence of ROP or ROP treatment not required for active ROP, i.e., presence of ROP is marked as "No" or the presence of active ROP requiring treatment is marked as "No".

5.5.3. Analysis of Additional Efficacy Variables

The following additional efficacy variables will be summarized descriptively by frequency tables.

- Proportion of patients with completion of vascularization of the peripheral retina to within 1 disc diameter of the ora serrata on FP at week 52 of chronological age
  - It will be summarized per eye and per patient.
- Presence of retinal vascularization and leakage on fluorescein angiography (FA) at week 52 of chronological age in patients participating in the FA substudy
  - It will be summarized per eye.

- Proportion of patients with regression of plus disease, regression of pre-retinal vascularized ridge, and progression of retinal vascularization beyond the ridge from baseline to week 52 of chronological age
  - It will be summarized per eye.
- Proportion of patients with progression to stage 4 or 5 ROP from baseline to week 52 of chronological age
  - It will be summarized per eye and per patient.
- Evaluation of visual function at week 52 of chronological age
  - It will be summarized per eye and per patient.
- Proportion of patients with requirement for sedation or general anesthesia to complete laser or aflibercept injection
  - It will be summarized per patient. The number of episode(s) will also be summarized.

Visual function will be presented by descriptive statistics for the following variables:

- fixation (central, steady and maintained) – Yes/No
- fixing and following a 5-cm toy – Yes/No
- in case visual function cannot be assessed with fixation, visual evoked potentials can be used for instance to evaluate if the visual function is normal or abnormal – Normal/Abnormal
- cycloplegic refraction (sphere, cylinder, axis, pseudophakia or intraocular vision correction) - Spherical equivalent
- ocular motility tests (if motility test done: assessment of ocular palsy with affected nerves, assessment of strabismus, and nystagmus). – Normal/Abnormal

The following time to event variables will be analyzed by patient level using the Kaplan-Meier estimates and displayed descriptively by treatment group. For patients with both eyes enrolled in the study, the time to event is defined as the time from randomization to the first event occurred in either eye.

- Time to recurrence of ROP
- Time to intervention with a second treatment modality for ROP or to development of unfavorable structural outcomes
- Time to completion of retinal vascularization
- Time required to perform aflibercept injection or initial laser treatment

5.5.4. Control of Multiplicity

To control the overall Type I error at α=0.05 (two-sided), an adjustment of 0.001 will be made for the DMC safety assessments. Therefore, at the final analysis, all nominal p-values will be compared to α = 0.049 and two-sided 95.1% confidence intervals will be reported.

If non-inferiority of the primary endpoint is declared significant, a hierarchical procedure for testing superiority will be used for the analysis of the secondary endpoints to control the overall alpha error rate at the 0.05 level based on the following order.

- Proportion of patients requiring intervention with a second treatment modality from baseline to week 52 of chronological age
- Proportion of patients with recurrence of ROP through week 52 of chronological age

The superiority of the primary endpoint will be tested, after testing for superiority of both secondary efficacy endpoints.

5.5.5. Subgroup Analyses

Subgroup analyses will be performed on the FAS population using descriptive statistics for primary efficacy based on the ROP status subgroup.

5.6. Analysis of Safety Data

The safety variables as described in Section 4.6 will be analyzed on SAF population. An additional pooled analysis may be performed between this study and studies #20090 (FIREFLEYE) and #20275 (FIREFLEYE NEXT) being conducted outside of the US by Regeneron's development partner, Bayer.

5.6.1. Adverse Events

TEAE summaries will be constructed displaying frequencies and proportions of patients reporting TEAEs within each SOC in decreasing order of total frequency according to the numbers of patients reporting the SOC and the preferred term (PT) within the SOC (not number of reports). The following categories will be summarized:

- Ocular TEAEs in the study eye(s)
- Ocular TEAEs in the fellow eye
- Non-ocular TEAEs

For the ocular TEAEs in the study eye(s), the frequency of the events and the number of study eyes(s) will be displayed as well. For the ocular TEAEs in the fellow eye, the number of patients with one study eye enrolled in each treatment group will be used as denominator.

Ocular TEAEs in the study eye(s) and non-ocular TEAEs will be further summarized for bilaterally and unilaterally treated patients separately.

Serious TEAEs, treatment-related TEAEs, treatment-related Serious TEAEs, and TEAEs leading to discontinuation will be summarized in the same way as described for TEAE.

TEAEs in the study eye related to the study conduct and those related to the study treatment will be summarized separately.

An overall summary of the AE profile for aflibercept in this study will be provided. A listing will be constructed that includes the patient identification, the treatment group, category of

TEAE (ocular study eye(s) or fellow eye, non-ocular), AE, MedDRA term, seriousness, severity, causality, elapsed time to onset, duration, and outcome.

Beyond TEAEs, All AEs from first dose of study treatment to end of study will be summarized in the same way as described for TEAE.

Subgroup analyses in TEAEs will be performed for the ROP status subgroup, for each of the following types of TEAE:

Summaries (by SOC and PT) of patients with:

- Ocular TEAEs study eye(s)
- Non-ocular TEAEs
- Serious ocular TEAEs study eye(s)
- Serious non-ocular TEAEs

5.6.2. Surgeries

An overall summary of number of patients undergoing surgery as described in Section 4.6.2 will be given by treatment group.

5.6.3. Clinical Laboratory

Clinical laboratory data will be collected only at screening. Data listings will be provided for hematology, chemistry and urinalysis data.

5.6.4. Vital Signs

Baseline vital signs and change from Baseline for vital sign variables described in Section 4.6.4 at each scheduled assessment visit will be summarized using descriptive statistics. These will include the number of patients, mean, median, standard deviation, minimum, and maximum.

5.6.5. Ocular Safety Measures

Baseline IOP and change from Baseline in IOP to each scheduled assessment visit will be summarized with descriptive statistics for study eye(s) treated with Aflibercept. Assessment of significant values or increases will be made and summarized for the proportion of patients with increased IOP in the study eye(s) with the categories defined in Section 4.6.5.

5.7. Analysis of Drug Concentration Data

The plasma concentrations of free and bound aflibercept over time will be summarized by descriptive statistics. No formal statistical hypothesis testing will be performed.

5.8. Pharmacokinetics and Pharmacokinetics/Pharmacodynamics Analyses

Population PK and exposure-response analyses for blood pressure may be performed, as appropriate.

5.9. Analysis of Immunogenicity Data

5.9.1. Analysis of ADA Data

The immunogenicity variables described in Section 4.7 will be summarized using descriptive statistics. Listings of ADA positivity and titers presented by patient, time point, and study treatment received will be provided. Incidence of treatment-emergent and treatment-boosted ADA will be assessed as absolute occurrence (N) and percent of patients (%), grouped by study treatment received.

Immunogenicity will be characterized by ADA responses and titers observed in patients in the ADA analysis set. ADA response categories and titer categories are defined as follows:

- Pre-existing immunoreactivity, defined as either an ADA positive response in the aflibercept ADA assay at baseline with all post dose ADA results negative, OR a positive response at baseline with all post first dose ADA responses less than 4 -fold over baseline titer levels.
- Treatment-emergent response, defined as a positive response in ADA assay post first dose when baseline results are negative or missing.
- Treatment-boosted response, defined as a positive response in the aflibercept ADA assay post first dose that is greater than or equal to 4-fold over baseline titer levels, when baseline results are positive.

Titer categories (Maximum titer values)

- Low (titer <1,000)
- Moderate (1,000 ≤ titer ≤ 10,000)
- High (titer >10,000)

6. DATA CONVENTIONS

The following analysis conventions will be used in the statistical analysis.

6.1. Definition of Baseline

Unless otherwise specified, the Baseline assessment for all measurements will be the last available valid measurement taken prior to the administration of investigational product.

6.2. Unscheduled Assessments

Assessments taken outside of protocol allowable windows will be displayed according to the CRF assessment recorded by the investigator.

Unscheduled assessments will not be included in the summaries.

If more than one value is available for a given visit, the value for this visit actually used for statistical summaries and analyses is as follows:

- The last non-missing repeated measurement, if the respective visit is before the start of treatment
- The first non-missing repeated measurement, if the respective visit is after the start of treatment

If an early termination visit is performed within the window (±1 week) of the next scheduled visit after the last visit, the efficacy assessment will be re-slotted to the next scheduled visit.

6.3. Handling of Patients who Discontinue

Patients who discontinue this study will not be replaced. The details for the handling of missing data due to patients who discontinue the study and study treatment are described in Section 6.4.1.

6.4. Handling of Missing Data

6.4.1. General Rules

When appropriate, the following rules will be implemented so as not to exclude patients from statistical analyses due to missing or incomplete data:

- Efficacy Variables

  For the primary, secondary and additional efficacy variables, missing observations will be imputed using LOCF if data are available from week 40 of chronological age visit. Otherwise, missing data will be considered as non-responders. The details are described in the efficacy analysis section (see Section 5.5).

- AE variables

  For some AEs it is important to determine whether the AE started before or after the first active aflibercept injection. If the AE start date is partially missing, it will be imputed by the latest possible date (considering other available data, e.g., stop date) to be conservative.

- Prior/concomitant medication

  For the tabulation of prior and concomitant medication, partially missing start dates of the medication will be imputed by the earliest possible time point, partially missing stop dates will be imputed by the latest possible time point.

7. INTERIM ANALYSIS

No formal interim analysis is planned. An independent data monitoring committee (DMC) will review the safety to determine if the study shows unacceptable risks for the patients, and to discontinue the study if necessary. Detailed information regarding the DMC procedures will be explained in a separate DMC charter.

An alpha adjustment of 0.001 will be made to the overall alpha at the final analysis in order to account for the DMC safety assessments.

8. SOFTWARE

All analyses will be done using SAS Version 9.4 or higher.

9. REFERENCES

1. RAINBOW Study: RAnibizumab Compared With Laser Therapy for the Treatment of INfants BOrn Prematurely With Retinopathy of Prematurity (RAINBOW). https://clinicaltrials.gov/ct2/show/NCT02375971
2. Salman AG and Said AM. Structural, visual and refractive outcomes of intravitreal aflibercept injection in high-risk prethreshold type 1 retinopathy of prematurity. Ophth Res 2015;53:15-20.
3. Huang CY, Lien R, Wang NK, et al. Changes in systemic vascular endothelial growth factor levels after intravitreal injection of aflibercept in infants with retinopathy of prematurity. Graefes Arch Clin Exp Ophthalmol 2018;256, 479-487.
4. Sukgen EA and Kocluk Y. Comparison of clinical outcomes of intravitreal ranibizumab and aflibercept treatment for retinopathy of prematurity. Graefe's Archive for Clinical and Experimental Ophthalmology. Revised 2018 Apr 16. Available from: https://doi.org/10.1007/s00417-018-4168-5.
5. ICH. (1998, February 5). ICH Harmonized tripartite guideline: Statistical principles for clinical trials (E9). International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use.
6. Rothmann M, Li N, Chen G, et al: Design and analysis of non-inferiority mortality trials in oncology. Stat Med 2003, 22:239-264.
7. Rothmann MD, Tsou HH: On non-inferiority analysis based on delta-method confidence intervals. J Biopharm Stat 2003, 13:565-583.
8. United States Food and Drug Administration: Guidance for Industry Non-Inferiority Clinical Trials 2010.

10. APPENDIX

10.1. Schedule of Time and Events:

Table 1: Schedule of Events

Visit Number[1]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16 EOS	Retreatment or rescue treatment follow-up visits			
Visit Name[1]	SCR	BL	D1	W1	W2	W3	W4	W6	W8	W10	W12	W16	W20	W24	W40 CA	W52 CA	TRT Visit	D1 FUP	1W FUP	4W FUP
Visit Window	–	–	+2D	±3D	±3D	±3D	+7D	±7D	±7D	±7D	±7D	±7D	±7D	±7D	±10D	±10D	–	+2D	±3D	+7D
Informed consent	X																			
In/exclusion criteria	X	X																		
Enroll (TVRS/TWRS)		X																		
Demographic	X																			
Medical history[2]	X																			
Prior/concomitant medication[3]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
AEs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Body weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Head circumference and body length	X								X			X		X	X	X				
Hearing test	X[4]																			
CNS imaging	X[5]																			
Physical exam[6]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Vital signs[7]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Anterior segment examination[8]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Binocular indirect ophthalmoscopy[9]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Tonometry[10]		X															X			
Digital wide-field retinal imaging[11]		X		X	X	X	X				X			X		X	X		X	X
Fluorescein angiography (for optional sub-study)[12]		X														X				
Blood sample for ADA (aflibercept group only)[13]		X									X									
Blood sample for PK (aflibercept group only)[14]			X		X		X													

Visit Number[1]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16 EOS	Retreatment or rescue treatment follow-up visits			
Visit Name[1]	SCR	BL	D1	W1	W2	W3	W4	W6	W8	W10	W12	W16	W20	W24	W40 CA	W52 CA	TRT Visit	D1 FUP	1W FUP	4W FUP
Visit Window	–	–	+2D	±3D	±3D	±3D	+7D	±7D	±7D	±7D	±7D	±7D	±7D	±7D	±10D	±10D	–	+2D	±3D	+7D
Hematology, clinical chemistry, urinalysis[15]	X																			
Urine protein test[16]		X	X		X		X													X
Aflibercept injection (for patients randomized to active)[17]		X															X			
Laser treatment (for patients randomized to control)[18]		X															X			
Visual function														X		X				
Refraction																X				
Ocular extrinsic motility														X		X				

Abbreviations CA = chronological age

1. Additional visits may be performed depending on the investigator's assessment of the patient's response to treatment, as part of local standard of care. Visits 1 and 2 can be on same day or within 10 days of each other. If treatment cannot take place at visit 2, it can be administered within 3 days of visit 2. For Visits 3 to 14, the intervals are based on the date of initial treatment. The intervals of the rescue/retreatment follow-up visits are based on the date of rescue/retreatment. A retreatment/rescue follow-up visit may be combined with a regularly scheduled visit. If a patient's second eye qualifies for treatment during the study, both eyes can follow the visit schedule of the first eye after the retreatment/rescue treatment follow-up visit schedule is completed for the second eye.
2. Maternal and patient medical history will be recorded.
3. Includes recording of oxygen supplementation. At screening, includes any of the medications listed in Section 8.7 given to mother during pregnancy and breastfeeding.
4. Hearing test to be performed once, at any time prior to discharge from neonatal intensive care unit (NICU).
5. Imaging is not required at screening if results from an imaging exam performed within the previous 10 days are available and there are no new neurological sign/symptoms. A second CNS imaging is to be carried out prior to discharge from the NICU.
6. Includes cardiovascular, respiratory, gastrointestinal, and neurological systems according to local general practice and aiming to evaluate overall health.
7. Temperature, blood pressure (before study treatment, if applicable), respiratory, and pulse rate.
8. Can be done using indirect ophthalmoscopy or portable slit lamp.
9. The pupils must be sufficiently dilated to allow examination of all ROP features. Unfavorable structural outcome is defined as retinal detachment, macular dragging, macular fold, or retrolental opacity.
10. Only in patients receiving aflibercept. The IOP will be measured in both eyes prior to the injection, and at least once post-injection (only in treated eyes).
11. Baseline imaging will be done prior to study treatment on the treatment day or up to 2 days prior to treatment. Imaging at week 24 and/or week 52 of chronological age may not be required if the following conditions are met: the reading center assessment of morphological outcomes on the week 12 image is in agreement with indirect ophthalmoscopy performed by the investigator, and the patient shows poor tolerance for the imaging procedure, preventing acquisition of images with adequate quality. Cases of rescue/retreatment require additional imaging before any treatment (on the treatment day or up to 2 days prior to treatment) as well as 1 week and 4 weeks following rescue/ retreatment.
12. Optional at baseline; mandatory at week 52 of chronological age (visit 16).

13. Only applicable for patients treated with aflibercept at baseline. One blood sample will be taken at baseline prior to dosing. At week 12, another blood sample will be taken prior to dosing to detect anti-drug antibodies (ADAs) and, if applicable, the occurrence of potential neutralizing antibodies (NAb) (if collection is not possible, the week 12 sample can be taken at week 16).
14. Only applicable for patients treated with aflibercept at baseline. One blood sample each for plasma concentrations of study drug will be taken on day 1 (approximately 24 hours after dosing), day 14, and day 28. Blood pressure must be measured before the PK sample is taken. If a patient's second eye is deemed eligible for treatment and is treated with aflibercept at visits 3, 5, or 7, the PK sample is taken before treatment.
15. A sample is not needed if results from laboratory tests within 8 days prior to screening are available and there was no change in the clinical situation from the time of the sample to the screening visit.
16. The baseline sample can be taken up to 2 days prior to baseline. The day 1 sample will be approximately 24 hours after study treatment.
17. The injection should be performed in both eyes on the same day, if applicable. After initial treatment, each study eye may receive up to 2 additional treatments (with a minimum interval of 28 days between injections) only if retreatment criteria are met. After retreatment, the retreatment/rescue treatment follow-up visits are also required. Blood pressure must be measured before the injection. Details of each injection will be recorded in the eCRF.
18. Multiple sessions performed within 1 week after baseline to complete the procedure will be counted as a single treatment. Each study eye may receive retreatment only if the retreatment criteria are met. After retreatment, the retreatment/rescue treatment follow-up visits are also required. Blood pressure must be measured before laser.

10.2. Summary of Statistical Efficacy Analyses

Endpoint	Major Analysis, population	Statistical Analysis	Sensitivity Analysis
Proportion of patients with absence of active ROP and of unfavorable structural outcomes at week 52 of chronological age	LOCF, FAS, investigator assessment	Non-inferiority of aflibercept vs. laser using CMH test adjusted by baseline ROP status	OC, PPS, RC data
Proportion of patients requiring intervention with a second treatment modality from baseline to week 52 of chronological age	LOCF, FAS	Superiority of aflibercept vs. laser using CMH test adjusted by baseline ROP status	
Proportion of patients with recurrence of ROP through week 52 of chronological age	LOCF, FAS	Superiority of aflibercept vs. laser using CMH test adjusted by baseline ROP status	
Time to recurrence of ROP	FAS	Kaplan-Meier estimate	
Time to intervention with a second treatment modality for ROP or to development of unfavorable structural outcomes	FAS	Kaplan-Meier estimate	
Completion of vascularization of the peripheral retina to within 1 disc diameter of the ora serrata on FP at week 52 of chronological age	FAS	Descriptive Statistics	
Assessment of retinal vascularization and leakage on fluorescein angiography (FA) at week 52 of chronological age in patients participating in the FA substudy	FAS	Descriptive Statistics	
Time to completion of retinal vascularization	FAS	Kaplan-Meier estimate	
Regression of plus disease, regression of pre-retinal vascularized ridge, and progression of retinal vascularization beyond the ridge from baseline to week 52 of chronological age	FAS	Descriptive Statistics	
Progression to stage 4 or 5 ROP from baseline to week 52 of chronological age	FAS	Descriptive Statistics	
Evaluation of visual function at week 52 of chronological age	FAS	Descriptive Statistics	
Time required to perform aflibercept injection or initial laser treatment	FAS	Kaplan-Meier estimate	
Requirement for sedation or general anesthesia to complete laser or aflibercept injection	FAS	Descriptive Statistics	

10.3. Calculation of confidence intervals using Mantel-Haenszel weighting scheme

The confidence intervals using the Mantel-Haenszel weighting scheme will be calculated according to the formulas given by Koch et al. (1990, p. 415 ff.)[2], i.e. to compute confidence intervals for the difference in two binomial proportions obtained from a multicenter trial, we calculate a weighted difference and its associated variance using Mantel-Haenszel weighting scheme.

For a multicenter study with h 2x2 tables, the weighted difference is:

$$d = \left(\sum w_h(p_{he} - p_{hs})\right)/\left(\sum w_h\right)$$

where $w_h = n_{he}n_{hs}/(n_{he}+n_{hs})$

and p_{he} = success rate for experimental treatment in stratum h

p_{hs} = success rate for standard treatment in stratum h

n_{he} = number of patients under experimental treatment in stratum h

n_{hs} = number of patients under standard treatment in stratum h

The variance of the weighted difference is:

$$var(d) = \left(\sum w_h^2(p_{hs}(1-p_{hs})/(n_{hs}-1) + p_{he}(1-p_{he})/(n_{he}-1))\right)/\left(\sum w_h\right)^2$$

A large sample approximation is used to compute the confidence interval:

$$CI = d \pm z_{\alpha/2}\, SQRT(var(d))$$

Where z_α is the α quantile of the standard normal distribution and SQRT is the square root function.

Signature Page for VV-RIM-00152939 v1.0

ESig Approval	[redacted]000
ESig Approval	[redacted]000
ESig Approval	[redacted]021 11:46:44 GMT+0000
ESig Approval	[redacted]-2021 20:56:45 GMT+0000
ESig Approval	[redacted]-2021 11:47:32 GMT+0000

Signature Page for VV-RIM-00152939 v1.0 Approved