DRCR Retina Network

Randomized Trial of Intravitreous Aflibercept Versus Intravitreous Bevacizumab + Deferred Aflibercept for Treatment of Center-Involved Diabetic Macular Edema (Protocol AC)

Statistical Analysis Plan

VERSION NUMBER	AUTHOR	APPROVER	EFFECTIVE DATE	REVISION DESCRIPTION
1.0	Wesley Beaulieu	Michele Melia	02 Feb 2019	Initial version for Protocol version 2.0
1.1	Wesley Beaulieu	Michele Melia	24 Apr 2020	Increased upper limit of 2-year visit window due to COVID-19 pandemic. Implemented strategies for controlling the Type I error rate. Modified diabetic retinopathy progression scale per specifications from fundus photograph reading center. Made clarifications. Revisions were made prior to initial data analysis. Applies to Protocol version 3.0
1.2	Wesley Beaulieu Danni Liu	Maureen Maguire	28 July 2021	Added methods to control for multiplicity. Other minor changes. Revisions were made prior to initial data analysis. Applies to Protocol version 4.0.
1.2	Danni Liu	Maureen Maguire	24 March 2022	Minor changes and clarifications to Sections 3.3 and 4.0.

SIGNATURES	
AUTHOR	
APPROVER	
PROTOCOL DIRECTOR	

1.0 Introduction

This document outlines the statistical analysis plan for the DRCR Retina Network protocol comparing the anti-vascular endothelial growth factor (anti-VEGF) drug aflibercept versus bevacizumab + deferred aflibercept for the treatment of center-involved diabetic macular edema (CI-DME) in eyes with at least moderate vision loss (Protocol AC). Moderate vision loss is defined as an Electronic-Early Treatment Diabetic Retinopathy Study (E-ETDRS) letter score of 68 to 24, which corresponds to an approximate Snellen equivalent of 20/50 to 20/320.

The primary objective of the protocol is to determine whether there is a treatment group difference in mean change in visual acuity from baseline over two years (area under the curve [AUC]). Participants will have visits every 4 weeks in year 1 and every 4 to 16 weeks in year 2, depending on the clinical course.

Study eyes will be assigned randomly to the two treatment groups in a 1:1 ratio. Participants may have one or two study eyes. Randomization of participants with one study eye will be stratified by site. Participants with two study eyes will have one eye randomized to each treatment, stratified by visual acuity of the eye with better visual acuity. If visual acuity is the same in both eyes, then the right eye will be considered to have better visual acuity than the left eye.

2.0 Primary Outcome Analysis

The primary analysis will consist of a treatment group comparison of mean change in visual acuity from baseline over two years adjusting for baseline visual acuity and the number of study eyes for the participant (one or two). Correlation arising from participants contributing two eyes to the analysis will be modeled by including a random intercept term for participant using a linear mixed effects model with robust variance estimation. The mathematical form of this model for eye *i* from participant *j* is as follows:

$$AUC_{ij} = \beta_0 + \beta_1 \times BaselineVision_{ij} + \beta_2 \times NumEyes_j + \beta_3 \times Treatment_{ij} + u_j + \epsilon_{ij}$$

$$u_j \sim N(0, \sigma^2_{participant})$$

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

The primary analysis is an intention-to-treat analysis. It will include all randomized eyes according to treatment group assignment at baseline. For each eye, AUC will be calculated by the trapezoidal rule using the following formula:

$$AUC = \sum_{i=1}^{n} \left(\frac{V_i + V_{i+1}}{2} \times (d_{i+1} - d_i) \right)$$

Where V_i is the visual acuity measured at the i^{th} visit, d_i is the number of days from randomization to the i^{th} visit, and n is the number of outcome visits included in the analysis. For presentation, AUC will be divided by the number of days between baseline and the n^{th} visit so

that the value shown will have units of letters rather than letter·days (e.g., 728 days for the primary outcome at 104 weeks). This statistic can then be interpreted as the average change in visual acuity over the time between baseline and the n^{th} visit. All analysis visits will be included for calculation of AUC and the number of days will be calculated based on the target date for the visit: 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 68, 84, and 104 weeks (see Section 8.2 for a description of analysis windows).

The *P* value, adjusted difference, and associated 95% confidence interval will be reported for the treatment group effect. If the *P* value for the test of the treatment effect is less than or equal to .05, then it will be concluded that there is a significant difference for change in visual acuity over two years between the groups. In other words, if $P \leq .05$, then the null hypothesis of $\beta_3 = 0$ will be rejected.

Markov Chain Monte Carlo (MCMC) multiple imputation with 100 imputations will be used to impute missing data. The imputation will be performed separately for each treatment group and the imputation model will include baseline visual acuity, change in visual acuity from baseline at all analysis visits, and the number of study eyes for the participant. Any imputed value outside of the range for electronic visual acuity measurements (<0 or >100 letters) will be truncated at the closest measurable value (i.e., 0 or 100).

A plot showing the change in visual acuity from baseline by treatment group over time will be constructed using observed data. In general, summary statistics (e.g., within-group means and standard deviations), will be based on observed data while estimates from statistical models (e.g., treatment group differences, confidence intervals, and *P* values) will be based on imputed data.

2.1 Sensitivity Analyses

Complete Case Analysis

A sensitivity analysis including only observed data from participants completing the 104-week visit (no imputation) will be conducted. If the analyses of imputed and observed data differ substantially, then exploratory analyses will be performed to evaluate factors that may have contributed to the difference. The sensitivity analysis of completers will only be performed if more than 10% of randomized participants do not complete the 104-week visit.

Tipping Point Analysis

Multiple imputation assumes that data are missing at random (MAR). In the present study, this would mean that whether change in visual acuity is missing or observed may be a function of observed baseline characteristics included in the imputation model, but not a function of the unobserved follow-up data being imputed. This assumption cannot be tested directly since these data are unknown. However, a tipping point analysis will be conducted to adjust the imputed values using a shift parameter and thereby determine how severe the departure from MAR must be to change the outcome of the analysis with respect to rejecting or failing to reject the null hypothesis.

A shift parameter will be applied to the imputed values in the aflibercept group to determine the tipping point at which the results of the primary analysis are nullified. That is, if one group is found to be superior ($P \leq .05$), the tipping point will identify the shift parameter necessary to yield $P > .05$. Conversely, if the null hypothesis is not rejected ($P > .05$), two tipping points will be identified – one that would make aflibercept superior and one that would make bevacizumab + deferred aflibercept superior. In either case, this tipping point will be evaluated to determine if it is plausible. If not, then the MAR assumption is likely reasonable. For example, if the tipping point were 100 letters, then this would be evidence that the MAR assumption is reasonable.

Per-Protocol Analysis

A per-protocol analysis will be conducted to estimate the treatment effect for each group among those not receiving any alternative treatment for DME (e.g., intravitreal corticosteroids). This analysis will include observed data from all randomized up to the time of alternative treatment for DME. Data collected after the alternative treatment will be set to missing prior to imputation. Imputation will otherwise be similar to the primary analysis. Note that focal/grid laser is allowed only if failure criteria have been met. The intention-to-treat analysis is considered the primary analysis. If the results of the primary and per-protocol analyses differ substantially, then exploratory analyses will be performed to evaluate factors that may have contributed to the difference. The per-protocol analysis will only be performed if more than 10% of randomized participants would be excluded by these criteria.

Confounding Analysis

Imbalances between groups in important covariates are not expected to be of sufficient magnitude to produce confounding in the primary analysis. However, the presence of confounding in the primary analysis will be evaluated in additional regression models using observed data (no imputation) by including baseline participant and study eye covariates including but not limited to the following:

- Age
- Duration of diabetes
- HbA1c
- Mean arterial blood pressure
- Prior PRP
- Prior treatment for DME
- Diabetic retinopathy severity level on fundus photographs as graded by the reading center
- Each of the following within 500 μm of the center of the macula on OCT as graded by the reading center (minimum 20 eyes in the cohort with and without the characteristic):
    - Epiretinal membrane
    - Vitreomacular traction
    - Cystoid abnormalities
    - Subretinal fluid

- Each of the following within 1800 μm of the center of the macula on fundus photography as graded by the reading center (minimum 20 eyes in the cohort with and without the characteristic):
    - Hemorrhages/microaneurysms
    - Hard exudates
    - Surface wrinkling retinopathy

Additional variables associated with the outcome will be included in regression models if there is an imbalance in the variables between treatment groups. Imbalance by treatment group will not be judged using statistical testing. Instead, imbalance will be judged by whether the size of the imbalance is clinically important, i.e., whether the imbalance is large enough to have a clinically important effect on the primary outcome.

2.2 Subgroup Analyses

Pre-planned subgroup analyses will repeat the primary analysis while including an interaction term for the baseline subgroup factor by treatment. Missing data will be imputed by treatment group similarly to the primary analysis except that the subgroup factor will be included in the imputation model.[1]

A significant ($P \leq .05$) type III test of the interaction term will be taken as an indication that subgroup effects need to be explored for full interpretation of the trial results. It is recognized that the study is not powered to detect subgroup effects and that lack of significance for the subgroup tests of interaction is not necessarily an indication that subgroup effects do not exist.

Interpretation of subgroup analyses will depend on whether the overall analysis demonstrates a significant treatment effect. In the absence of a significant treatment effect in the primary analysis, subgroup analyses will be interpreted with caution.

Baseline variables to be evaluated for subgroup effects include the following:

- OCT central subfield thickness: < 400 μm vs ≥ 400 μm (Stratus equivalent)
- Visual acuity: 20/50 to 20/63 vs. 20/80 to 20/320

Subgroups will only be analyzed if there are at least 20 eyes in each treatment group for each subgroup to increase statistical precision. Cutoffs of continuous outcomes may be modified to achieve a reasonable number of eyes in each group.

The above subgroups are considered of primary interest. For each variable, the rationale for performing the analysis is listed in Table 1 below.

Table 1. Subgroup analyses.

Variable	Rationale
OCT central subfield thickness	Eyes with greater OCT central subfield thickness may have higher anti-VEGF levels, leading to a larger relative treatment effect for the drug with stronger VEGF binding affinity (aflibercept).
Baseline visual acuity	Eyes with lower visual acuity may have higher VEGF levels, leading to a larger relative treatment effect for the drug with stronger VEGF binding affinity (aflibercept).

The following subgroup factors will be evaluated in exploratory analyses. Only point estimates and 95% confidence intervals for the within-subgroup treatment effects will be presented – *P* values will not be presented. The finding of a subgroup effect for any of these factors will be interpreted as hypothesis generating only and in need of confirmation from further studies.

- Prior treatment for DME: yes vs. no
- HbA1c: $< 7.5\%$ vs. $\geq 7.5\%$
- Sex: female vs. male
- Race/Ethnicity: Non-Hispanic White vs. Non-Hispanic Black/African American vs. Hispanic or Latino (exclude all other groups due to anticipated small sample size)

2.3 Center Effects

The number of study participants per center is expected to be small for many centers. Therefore, center effects will not be included in the statistical model. However, for centers with a large number of study participants ($N \geq 20$ in either treatment group), heterogeneity across centers will be explored using random center effects by estimating empirical best linear unbiased predictors along with 95% confidence intervals.

3.0 Secondary Outcome Analyses

3.1 Visual Acuity and OCT

Additional analyses of visual acuity will use the imputed data sets created for the primary outcome. For OCT outcomes, new imputed data sets will be created similarly by substituting baseline central subfield thickness (converted to a common scale based on the most accurate conversion algorithms available) and change in central subfield thickness from baseline for baseline visual acuity and change in visual acuity from baseline, respectively. A plot of mean change in OCT central subfield thickness over time by group will be constructed using observed data.

Analyses will be conducted at 24, 52 and 104 weeks unless otherwise noted. P values will be calculated for mean change in visual acuity and mean change in OCT central subfield thickness at 104 weeks only. All other secondary outcomes will be summarized with a model-based point estimate and 95% between-group confidence interval (no P value).

Analyses of continuous outcomes will be conducted in a manner similar to the primary analysis using linear mixed effects models and substituting central subfield thickness for visual acuity, depending on the outcome.

Analyses of binary outcomes will be conducted using binomial regression with an identity link (estimating risk difference),[2] robust variance estimation, and adjustment for baseline visual acuity or OCT central subfield thickness (depending on the outcome). Generalized estimating equations (GEE) will be used to control for correlations arising from participants contributing two study eyes to the analysis. Baseline visual acuity (for visual acuity outcomes) or central subfield thickness (for central subfield thickness outcomes) and the number of study eyes will be included as covariates. The proportion of eyes meeting the outcome at the visit will be reported for each treatment group. In addition, the between-group risk difference and 95% confidence interval for the treatment effect will be presented; the *P* value will not be presented. If binomial regression fails to converge in one or more outcomes, then logistic regression with a random intercept for participant, conditional standardization, and the delta method (to estimate the risk difference)[3] may be used instead for all binary outcomes.

Table 2. Analyses of Secondary Visual Acuity and OCT Outcomes.

Outcome	Analysis Technique
Mean change in visual acuity from baseline	Linear mixed model [a]
Success proportion: visual acuity gain ≥ 15 letters	Binomial regression with GEE [b]
Success proportion: visual acuity gain ≥ 10 letters	Binomial regression with GEE [b]
Failure proportion: visual acuity loss ≥ 10 letters	Binomial regression with GEE [b]
Failure proportion: visual acuity loss ≥ 15 letters	Binomial regression with GEE [b]
Success proportion: visual acuity ≥ 84 letters (~20/20)	Binomial regression with GEE [b]
Success proportion: visual acuity ≥ 69 letters (~20/40)	Binomial regression with GEE [b]
Failure proportion: visual acuity ≤ 38 letters (~20/200)	Binomial regression with GEE [b]
Mean change in OCT central subfield thickness from baseline	Linear mixed model [a]
Proportion of eyes with OCT central subfield thickness below the gender-specific spectral domain OCT threshold for CI-DME [c]	Binomial regression with GEE [b]
Mean change in OCT retinal volume from baseline	Linear mixed model [b]

[a] At 24 and 52 weeks, a point estimate and 95% CI will be provided, but a *P* value will not be provided. At 104 weeks, a *P* value will be provided only if the primary analysis demonstrates a significant difference.
[b] A point estimate and 95% CI will be provided, but a *P* value will not be provided.
[c] For Zeiss Cirrus, ≥ 290 µm for women and ≥ 305 µm for men. For Heidelberg Spectralis, ≥ 305 µm for women and ≥ 320 µm for men. No imputation for this outcome because the threshold values are machine specific and not Stratus equivalents, which are being imputed.

3.2 Changes in Diabetic Retinopathy

The proportion of eyes with 2-step improvement or worsening of diabetic retinopathy on fundus photographs (defined in Table 3) will be assessed at 52 and 104 weeks using observed data only (no imputation). Analyses will be conducted with binomial regression while adjusting for baseline diabetic retinopathy severity (ordinal transformation) and using GEE to control for the correlation arising from participants contributing two study eyes to the analysis. Baseline ordinal diabetic retinopathy severity and the number of study eyes will be included as covariates. The proportion of eyes meeting the outcome at the visit will be reported for each treatment group. In addition, the risk difference and 95% confidence interval for the treatment effect will be reported, but not a *P* value.

Table 3. Definitions of Diabetic Retinopathy Improvement and Worsening for Eyes

Baseline		Worsening (if follow up ≥)	Improvement (if follow up ≤)
NPDR	10/12	35	*Exclude*
	14/15/20	43	*Exclude*
	35	47	10/12
	43	53	14/15/20
	47	61	35
	53	61	43
PDR	61	71	53
	65	75	53
	71	81	61
	75	81	65
	81	*Exclude*	71
	85	*Exclude*	75

Abbreviations: NPDR, non-proliferative diabetic retinopathy; PDR, proliferative diabetic retinopathy.

3.3 Additional Secondary Outcomes

Additional secondary outcomes will be evaluated 104 weeks, unless otherwise specified below.

Table 4. Additional Secondary Outcomes.

Outcome	Analysis Technique [a]
Time to receipt of panretinal photocoagulation, vitrectomy, or occurrence of vitreous hemorrhage, traction retinal detachment, neovascularization of the iris, or neovascular glaucoma [b]	Marginal Cox proportional hazards regression
Number of injections up to the visit [c]	Linear mixed model (52 and 104 weeks)
Number of visits [d]	Independent samples *t*-test

[a] Point estimates and 95% confidence intervals will be presented, but *P* values will not be presented.
[b] Includes all randomized eyes, regardless of visit completion, and all visits up to and including the 104-week visit

[c] Limited to eyes that complete the visit or any later common visit

[d] Evaluated at 104 weeks for 104-week completers among participants with one study eye. Bilateral participants are excluded because number of visits is a participant-level outcome and treatment group is an eye-level variable.

Kaplan-Meier estimates for the time to receipt of panretinal photocoagulation, vitrectomy or occurrence of vitreous hemorrhage, traction retinal detachment, neovascularization of the iris, or neovascular glaucoma will be provided. A marginal Cox proportional hazards model with a robust sandwich estimate of the covariance matrix will be used to control for correlations arising from bilateral participants.[4] Baseline ordinal diabetic retinopathy severity and the number of study eyes will be included as covariates. The model will adjust for ordinal baseline diabetic retinopathy severity because these events all represent worsening of diabetic retinopathy, and eyes with more advanced diabetic retinopathy at baseline are more likely to experience worsening. The treatment effect will be measured as a hazard ratio and will be reported with the associated 95% confidence interval, but not a *P* value. When the number of events is low (i.e., less than 10 in either treatment group), these time-to-event analyses will be replaced by comparing the percentage in each treatment group with the PDR event at any time during follow-up. A 95% confidence interval on the difference in proportions will be calculated by inverting two separate one-sided exact tests that are based on the score statistic.

For comparison of the number of injections, a linear mixed model will be used with the number of study eyes included as a covariate. For comparison of the number of visits, an independent samples t-test limited to participants with one study eye will be used.

4.0 Outcomes within Treatment Groups

Within each treatment group, the following outcomes will be tabulated without formal statistical comparisons.

- Distribution and mean (standard deviation) number of intravitreous injections performed up to 24, 52, and 104 weeks as well as the intervening periods for eyes completing any common visit at or beyond the upper limit (e.g., for injections through 52 weeks, eyes must have completed the 52-, 68-, 84-, or 104-week visit).
    - Intervals will be closed on the left and open on the right (e.g., for injections through 24 weeks, an injection given at 24 weeks will not be counted towards the total; however, an injection given at 24 weeks will count for the interval of 24 to 52 weeks), unless the upper boundary is at 104 weeks in which case the interval is closed on the left and right.
- Proportion of eyes that met switch criteria by the 12, 24, 52, or 104-week visits.
    - Estimated using the Kaplan-Meier method.

5.0 Economic Analysis

The purpose of the economic analysis is to compare the treatment groups with respect to cost. An incremental cost effectiveness ratio (ICER) will be calculated. Data from the clinical trial on the

number of clinic visits completed, number of procedures performed (e.g., OCT, fundus photographs), and number of aflibercept and bevacizumab treatments will be used to estimate an average cost per patient for each treatment arm. The Medicare Fee Schedule will be used to estimate medical costs. For outcomes measured at the participant level, bilateral participants are non-informative with respect to the treatment comparison and will not be included in the economic analyses.

6.0 Safety Analysis

Adverse events will be categorized as study eye, fellow eye, or systemic. A full listing of adverse events will be tabulated by treatment. An additional tabulation will be made for adverse events possibly related to study treatment.

All randomized eyes will be included in the safety analysis and analyzed according to treatment assignment at randomization, regardless of treatment actually received. Any adverse event that occurred up to and including the 104-week visit (or, if the participant did not complete the 104-week visit, 896 days from randomization, the end of the 2-year analysis window) will be reported.

6.1 Study Eye Adverse Events

The frequency of the event occurring at least once will be calculated for study eyes in each treatment group. The proportion of eyes experiencing each outcome will be compared between treatment groups using Barnard's unconditional exact test, considering the number of eyes in each treatment group fixed. It is noted that this method does not adjust for the potential correlation arising from participants with two study eyes; however, given the low expected frequency of adverse events and small proportion of bilateral subjects, the impact should be minimal. Adjustment for such correlations would be statistically problematic due to the low frequency of events.

The following ocular adverse events will be assessed:

- Endophthalmitis
- Retinal detachment (rhegmatogenous, traction, combined rhegmatogenous and traction, or not otherwise specified)
    - Rhegmatogenous retinal detachment (tabulated without formal analysis)
    - Traction retinal detachment (tabulated without formal analysis)
- Traumatic cataract due to intravitreous injection
    - Analysis limited to eyes with phakic lens at baseline
- Vitreous hemorrhage
- Ocular inflammation

- Intraocular pressure (IOP) elevation (composite outcome; individual components below will be tabulated without formal analysis):
  - Increase in IOP ≥ 10 mmHg from baseline (at a follow-up visit)
  - IOP ≥ 30 mmHg (at a follow-up visit)
  - Initiation of medication to lower IOP that was not in use at baseline
  - Glaucoma procedure
- Neovascularization of the iris
- Neovascular glaucoma

6.2 Systemic Adverse Events

Systemic adverse events will be reported in three groups: (1) unilateral participants randomized to bevacizumab + deferred aflibercept, (2) unilateral participants randomized to aflibercept, and (3) bilateral participants randomized to bevacizumab + deferred aflibercept in one eye and aflibercept in the other eye. The frequency of each event occurring at least once per participant will be calculated. The proportion of participants with each systemic adverse event will be compared among groups using Fisher's exact test. If the overall test has $P \leq 0.05$, then pairwise comparisons between groups will be conducted using Fisher's exact test without the need to adjust for multiple comparisons.[5] The following systemic adverse events will be assessed:

- Primary:
  - Death
  - Serious adverse event
  - Hospitalization
  - Cardiovascular and cerebrovascular events according to the Antiplatelet Trialists' Collaboration (excerpted from BMJ Jan 8, 1994):
    - Nonfatal myocardial infarction
    - Nonfatal stroke (counted only if symptoms lasted at least 24 hours)
    - Death attributed to cardiac, cerebral, hemorrhagic, embolic, other vascular (does not need to be ischemic in origin), or unknown cause
    - At least one event (nonfatal myocardial infarction, nonfatal stroke, or death attributed to potential vascular or unknown cause)

Note that transient ischemic attack, angina, possible myocardial infarction, and possible stroke are not counted. Nonfatal myocardial infarction and nonfatal stroke require that the patient be alive at the end of the study. If not, then only the death is counted.

- Secondary (for tabulation without statistical comparison):
  - Frequency of at least one event per participant in each Medical Dictionary for Regulatory Activities (MedDRA) system organ class

7.0 Additional Tabulations

The following will be tabulated according to treatment group:

- Baseline demographic and clinical characteristics
- Visit completion rate for each annual visit (excluding deaths)
- Treatment adherence

8.0 General Principles for Analysis

8.1 Analysis Cohort

Unless otherwise stated, all treatment comparison analyses will follow the intention-to-treat principle with all randomized eyes included and each eye analyzed according to the randomized treatment assignment, regardless of treatment actually received.

8.2 Visit Windows for Analysis

For common visits, the analysis windows will be defined according to Table 5. Note that all eyes have visits every 4 weeks in year 1, but the protocol visit schedule varies in year 2 depending on the clinical course. Therefore, visit windows at 68 and 84 weeks have been defined for analysis purposes.

If multiple visits fall within the same analysis window, the following algorithm will be used to prioritize which visit will be used for analysis:

- Visits with non-missing visual acuity data will be prioritized over visits with missing data
- If there is no protocol visit in the analysis window, then the visit closest to the target will be used
  - For visits falling in more than one window, priority will be given to the 24-, 52-, and 104-week visits. Otherwise, the visit will be assigned to the earlier window.

Table 5. Analysis Windows for Outcome Visits

Protocol Visit	Target	Analysis Window	
4 weeks	28 days	14 – 42 days	(4 ± 2 weeks)
8 weeks	56 days	42 – 70 days	(8 ± 2 weeks)
12 weeks	84 days	70 – 98 days	(12 ± 2 weeks)
16 weeks	112 days	98 – 126 days	(16 ± 2 weeks)
20 weeks	140 days	126 – 154 days	(20 ± 2 weeks)
24 weeks	168 days	154 – 182 days	(24 ± 2 weeks)
28 weeks	196 days	182 – 210 days	(28 ± 2 weeks)
32 weeks	224 days	210 – 238 days	(32 ± 2 weeks)
36 weeks	252 days	238 – 266 days	(36 ±2 weeks)
40 weeks	280 days	266 – 294 days	(40 ± 2 weeks)
44 weeks	308 days	294 – 322 days	(44 ± 2 weeks)
48 weeks	336 days	322 – 350 days	(48 ± 2 weeks)
52 weeks	364 days	308 – 420 days	(52 ± 8 weeks)
68 weeks	476 days	420 – 532 days	(68 ± 8 weeks)
84 weeks	588 days	532 – 644 days	(84 ± 8 weeks)
104 weeks	728 days	644 – 896 days	(92–128 weeks)

8.3 Missing Data

In general, the strategy for handling missing data is included with the description of each individual analysis. Where not otherwise specified, only participants with non-missing data are included in the analysis.

8.4 Outliers

To help ensure that statistical outliers do not have undue impact on analyses of continuous visual acuity and OCT outcomes (including the primary outcome), change in visual acuity change in central subfield thickness, and change in retinal volume will be truncated to ± 3 standard deviations. The standard deviations will be based on observed data from 104-week completers at the 104-week visit, irrespective of treatment group. Truncation will be performed after imputation of missing data, where applicable (i.e., raw data will be used for imputation). For the primary outcome, AUC will be calculated based on the imputed values after truncation. There will be no truncation of the AUC outcome itself.

8.5 Model Assumptions

All model assumptions, including linearity, normality of residuals, heteroscedasticity, and proportional hazards will be verified. If model assumptions are not reasonably satisfied, then covariates may be categorized or excluded, and a nonparametric approach, transformation, or robust method may be considered.

8.6 Multiple Testing

The primary analysis will be conducted at alpha of 0.05. If the primary analysis demonstrates a significant treatment group difference, then mean change in visual acuity from baseline at 104 weeks and mean change in OCT central subfield thickness from baseline at 104 weeks will be tested as secondary outcomes. The Holm method will be used to provide strong control of alpha at 0.05.[5] If the primary analysis fails to show a significant difference, then outcomes will be described with summary statistics, model-based point estimates, and between-group 95% confidence intervals without a P value. This approach controls the family-wise error rate at 5%.

There will be no formal adjustment for multiplicity in sensitivity, subgroup, or safety analyses. For exploratory subgroup analyses, the number of significant results expected by chance given the number of comparisons will be noted.

9.0 Example SAS Code

Below is an example of SAS code for the primary outcome analysis. The actual code used might differ due to variable naming conventions.

```
/* Generate Imputed Data Sets */
proc mi data=studyEyes nimpute=100 seed=9999 out=outMI;
  var va0 vaChg4 vaChg8 vaChg12 vaChg16 vaChg20 vaChg24 vaChg28
      vaChg32 vaChg36 vaChg40 vaChg44 vaChg48 vaChg52 vaChg68
      vaChg84 vaChg104 switchGrpFlg bilateralFlg;
run;

/* ...DATA steps to truncate changes in visual acuity and calculate primary
outcome (vaChgAUC104)... */

/* Linear Mixed Model for Each Imputation */
proc mixed data=outMI empirical;
  by _imputation_;
  class PtID;
  model vaChgAUC104=va0 bilateralFlg switchGrpFlg / solution
    covb ddfm=kr2;
  random int / subject=PtID;
  ods output solutionF=mixParms covb=mixCovB;
run;
```

```
/* Combine Results */
proc mianalyze data=mixParms covb(effectvar=rowcol)=mixCovB
  edf=308;
/* 312 eyes minus 3 fixed effects + intercept = 308 degrees of freedom */
  modeleffects va0 bilateralFlg switchGrpFlg;
run;
```

10.0 References

1. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? *Stat Methods Med Res.* 2018;27(9):2610-2626.
2. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol.* 2005;162(3):199-200.
3. Localio AR, Margolis DJ, Berlin JA. Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. *J Clin Epidemiol.* 2007;60(9):874-882.
4. Lee EW, Wei LJ, Amato DA. *Cox-type regression analysis for large numbers of small groups of correlated failure time observations.* Dordrecht: Kluwer Academic; 1992.
5. Dmitrienko A, D'Agostino RB, Sr. Multiplicity Considerations in Clinical Trials. *N Engl J Med.* 2018;378(22):2115-2122.