Diabetic Retinopathy Clinical Research Network

Treatment for Central-Involved Diabetic Macular Edema in Eyes with Very Good Visual Acuity (Protocol V)

Statistical Analysis Plan

NCT01909791

VERSION NUMBER	AUTHOR	APPROVER	EFFECTIVE DATE	REVISION DESCRIPTION
1.0	Wesley T. Beaulieu	Michele Melia	12 April 2017	Initial version for Protocol version 3.0
2.0	Wesley T. Beaulieu		27 July 2018	Revisions following DSMC review of SAPs to harmonize with other DRCR.net SAPs. Still applies to Protocol version 3.0. Changes made after interim analysis but prior to primary analysis. Changes were not a result of the interim analysis. Section 9.0 outlines the changes made and the rational for each change.

1.0 Introduction

This document outlines the statistical analysis plan for the Diabetic Retinopathy Clinical Research Network (DRCR.net) Protocol V comparing prompt focal/grid photocoagulation with deferred intravitreous anti-vascular endothelial growth factor (VEGF), observation with deferred intravitreous anti-VEGF, and prompt intravitreous anti-VEGF for treatment of central-involved diabetic macular edema (CI-DME). The anti-VEGF agent used in this trial is aflibercept (Eylea®, Regeneron Pharmaceuticals, Tarrytown, NY).

The primary objective of the protocol is to determine if there are differences between the three treatment groups in the proportion of eyes that lose 5 or more letters of visual acuity from baseline at the 2-year (104-week) visit. Several secondary outcomes will be analyzed. These include the mean change in visual acuity, the proportion of eyes with 20/20 vision or better, and the mean change in OCT central subfield thickness (CST). Analyses with visual acuity or CST as outcomes (continuous or binary) will adjust for baseline visual acuity or CST, respectively, as the average of the two values from the screening and randomization visits. If the average results in a decimal (.5), the value will be rounded up if the randomization value is higher than the screening visit value, otherwise it will be rounded down so that the baseline value is always an integer. This rounding approach is the same method used in the protocol when calculating change in visual acuity for retreatment.

Study eyes are assigned randomly to the three treatment groups in a 1:1:1 ratio stratified by site and recent (within 4 months) or planned DME treatment in the fellow eye (yes or no). Participants may have only one study eye enrolled in the randomized trial. The fellow eye may, however, be enrolled in the observational phase. Eyes enrolled in the observational phase will be used to collect data on the natural history of diabetic macular edema and are used in exploratory analyses. For eyes entering the randomized trial from the observational phase, the last observational follow-up visit may be used in place of the screening visit.

2.0 Efficacy Analysis Plan

2.1 Primary Outcome Analysis

The primary outcome is a 5 or more letter decrease in visual acuity letter score from baseline at 104 weeks. The primary analysis will be an intention-to-treat analysis that includes all randomized eyes according to the treatment group assignment at randomization. Treatment group comparisons will be conducted using binomial regression with an identity link (estimating risk difference) and adjusting for baseline visual acuity and recent or planned DME treatment in the fellow eye at the time of randomization. If binomial regression is not feasible then Poisson regression with a robust error variance and log link (estimating risk ratio) will be used while still adjusting for baseline visual acuity and recent or planned DME treatment in the fellow eye at the time of randomization (Spiegelman and Hertzmark 2005). If Poisson regression is needed for one or more analyses, then Poisson regression may be used for analysis of all binary outcomes to harmonize the presentation of results.

The primary analysis will include three two-group comparisons of the proportion of eyes meeting the primary outcome at 104 weeks. Experiment-wise Type 1 error rate (α) will be 0.05 (2-sided). The Hochberg (1988) procedure will be used to control the overall Type 1 error for multiple comparisons. This procedure contrasts *P* values p(1), p(2), p(3), ordered from least to greatest, with a set of critical values and rejects all null hypotheses with smaller or equal *P* values to that of any one found less than its critical value. It can be summarized as follows:

- If $p(3) \leq \alpha$, then stop and reject $H_0(1)$, $H_0(2)$, $H_0(3)$ at level α; otherwise fail to reject $H_0(3)$ and go to the next step
- If $p(2) \leq \alpha/2$ then stop and reject $H_0(1)$ and $H_0(2)$; otherwise fail to reject $H_0(2)$ and go to the next step
- If $p(1) \leq \alpha/3$ then stop and reject $H_0(1)$; otherwise fail to reject $H_0(1)$

For the primary analysis, $\alpha = 0.05$, p(k) = kth highest ordered *P* value, $H_0(k)$ = hypothesis with *P* values p(k), k = 1, 2, 3

Multiple comparisons will be handled similarly for all secondary and exploratory outcomes (i.e., visual acuity, CST, and diabetic retinopathy severity).

Missing data will be imputed with Markov chain Monte Carlo (MCMC) multiple imputation. The imputation model will include treatment group, recent or planned DME treatment in the fellow eye, the level of baseline visual acuity in the study eye, and the change in study-eye visual acuity from baseline to all common follow-up visits (i.e., 8, 52, and 104 weeks). The primary outcome of a 5 or more letter loss in visual acuity from baseline to 104 weeks will be calculated from the imputed data sets. To reiterate, continuous missing visual acuity letter scores are what will be imputed using the MCMC model; the missing binary primary outcome will be calculated using the imputed letter scores.

A plot showing the primary outcome over time will be constructed using observed data. In general, summary statistics (e.g., proportion of eyes losing 5 or more letters from baseline) for all outcomes, will be based on observed data while numbers from statistical models (e.g., treatment group differences, confidence intervals, and *P* values) will be based on imputed data, where applicable.

2.1.1 Sensitivity Analyses

A sensitivity analysis that does not use multiple imputation and only includes observed data (i.e., only participants who complete the 104-week visit with no imputation of missing data) will be conducted. If the results from the primary analysis and sensitivity analysis are discrepant, an exploratory analysis will be carried out to identify factors that contributed to the difference.

In the event that outcome rates are much lower than expected, i.e. fewer than 5 events in one or more treatment groups, a second sensitivity analysis, also using observed data (no imputation), will be conducted using exact logistic regression adjusting for baseline visual acuity and recent

or planned DME treatment in the fellow eye. In this case, results from exact logistic regression may be more robust than results from binomial regression. If differences from the two analyses emerge, exploratory analyses will be carried out to identify factors that contributed to the difference.

2.1.2 Per-Protocol Analysis

A per-protocol analysis of the primary outcome will be conducted using only observed data (no imputation) in which any eye receiving a treatment for DME prior to 2 years other than focal/grid photocoagulation or an anti-VEGF injection will be excluded. If the results differ from the primary intention-to-treat analysis, exploratory analyses will be performed to identify factors that 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.

2.1.3 Confounding

Imbalances between groups in important covariates are not expected to be of sufficient magnitude to produce confounding. However, the presence of confounding will be evaluated in regression models using observed data by including baseline participant and study eye covariates including but not limited to the following: age, duration of diabetes, HbA1c, prior treatment for DME, and diabetic retinopathy severity as graded by the photograph reading center. 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.

2.1.4 Subgroup Analyses

Pre-planned subgroup analyses will repeat the primary analysis while including an interaction term for the baseline subgroup factor by treatment. Only observed data (no imputation) will be used for these analyses. Unless the imputation process is done separately for each treatment group and the subgroup factor is included in the imputation model, the analysis is biased towards the null hypothesis of no interaction when an interaction is present (Sullivan et al., 2016). It is recognized that analyzing only observed data also may be biased, but unlike the imputed analysis, it is not necessarily biased in the presence of interaction. In addition, the usual procedure for combining results across imputed datasets in not necessarily valid for the global test of interaction.

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. In addition, within-subgroup treatment effects and 95% confidence intervals will be estimated from the interaction model if the interaction P value is less than .05. 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, analyses of subgroups will be interpreted with caution.

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

- OCT central subfield thickness: continuous and < 400 µm vs. ≥ 400 µm
- Diabetic retinopathy severity level from fundus photographs: continuous (ordinal numeric transformation of retinopathy severity grades) and categorical (proliferative diabetic retinopathy [PDR] vs. non-proliferative diabetic retinopathy [NPDR] or less)
- Presence of central epiretinal membrane or vitreomacular traction graded on OCT by the central reading center: yes vs. no

Note that 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 and ordinal outcomes may be modified to achieve a reasonable number of eyes in each group.

The above subgroups are considered those of primary interest for which a rationale for a subgroup effect is hypothesized. For each factor, the rationale for performing the analysis is listed in Table 1 below.

Table 1. Subgroup analyses.

Factor	Rationale
OCT central subfield thickness	There are conflicting reports in the literature as to whether laser is relatively less effective when used in thicker retinas as compared to thinner ones when comparing to anti-VEGF (i.e., there is an quantitative interaction; anti-VEGF is expected to be better than laser for both thick and thin retinas, the difference will be greater for thick retinas).
Diabetic retinopathy severity level from fundus photographs	Eyes with more advanced retinopathy may have better outcomes with anti-VEGF, which is known to be effective in treating diabetic retinopathy.
Presence of epiretinal membrane or vitreomacular traction	Eyes with epiretinal membrane or vitreomacular traction have thickening that is less likely to resolve and may meet failure criteria faster and receive anti-VEGF. This could make outcomes from the observation and laser groups more similar to the anti-VEGF group because they will be treated more similarly.

The following subgroup factors also will be evaluated in exploratory analyses. The finding of a significant subgroup effect for any of these factors will be interpreted as hypothesis generating only and in need of confirmation from further studies.

- Leakage patterns identified on fluorescein angiography and clinical exam: yes vs. no
- Presence of circinate ring: yes vs. no
- Duration of diabetes: continuous and categorical (dichotomized based on a clinically-relevant cut point)
- Duration of DME: continuous and categorical (dichotomized based on a clinically-relevant cut point)
- Lens status: phakic vs. pseudophakic
- Prior DME treatment: yes vs. no
- Prior focal/grid laser for DME: yes vs. no
- Prior anti-VEGF for DME: yes vs. no
- Prior panretinal photocoagulation (PRP) treatment: yes vs. no
- Age: continuous and < 60 vs. ≥ 60 years
- HbA1c: continuous and $< 7.5\%$ vs. $\geq 7.5\%$
- Sex: female vs. male
- Race/Ethnicity: non-Hispanic white vs. black/African American vs. Hispanic (exclude all others due to anticipated small sample size) as well as white vs. non-white

Interaction *P* values will be calculated using the continuous or ordinal variables where possible in addition to the categorizations described above.

2.1.5 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 ($N \geq 20$ per treatment group), heterogeneity across centers will be explored using random center effects by estimating empirical best linear unbiased predictors along with 95% confidence intervals.

2.2 Secondary Analyses of Visual Acuity

Additional analyses will be conducted on the visual acuity data, primarily to aid clinicians and patients in the interpretation of the primary outcome results and to explore treatment group effects at other follow-up times (52 weeks). The secondary visual acuity outcomes and the analysis methods are specified in Table 2. All analyses will include adjustments for baseline visual acuity and recent or planned DME treatment in the fellow eye. With the exception of low-

contrast visual acuity, analyses will use the imputed data sets created for the calculation of the primary outcome. For low-contrast visual acuity, a new group of imputed data sets will be created as described in section 2.1 except low-contrast visual acuity will be substituted for visual acuity. Only eyes with low-contrast visual acuity measurements at baseline will be included since not all sites have low-contrast visual acuity capability.

Table 2. Additional Analyses of Visual Acuity.

Outcome	Analysis Technique
Failure proportion: Worsening ≥ 15 letters	Binomial regression
Failure proportion: Worsening ≥ 10 letters	Binomial regression
Failure proportion: Worsening ≥ 5 letters*	Binomial regression
Success proportion: Improvement ≥ 5 letters	Binomial regression
Proportion with study-eye visual acuity ≥ 84 letters (approximately 20/20)	Binomial regression
Mean change in visual acuity	ANCOVA
Mean change in low-contrast visual acuity†	ANCOVA

*At 52 weeks only. The primary analysis evaluates this response at 104 weeks.

†Adjusting for baseline low-contrast visual acuity instead of visual acuity.

A plot that shows the mean change in visual acuity from baseline by treatment group over time will be constructed with observed data.

2.3 Analysis of Retinal Thickness Secondary Outcomes

Several OCT CST outcomes are of interest and will be evaluated at 52 and 104 weeks. Each analysis will adjust for baseline CST and recent or planned DME treatment in the fellow eye. All CST values will be converted to time-domain equivalents prior to analysis. The outcomes and the analysis techniques to be used are specified in Table 3. Analyses will use multiply imputed data sets created as described in section 2.1 but substituting CST for visual acuity.

Table 3. Analyses of Retinal Thickness.

Outcome	Analysis Technique
Mean change in OCT CST	ANCOVA
Success proportion: 2 log step increase in CST	Binomial regression
Success proportion: 1 log step increase in CST	Binomial regression
Failure proportion: 1 log step decrease in CST	Binomial regression
Failure proportion: 2 log step decrease in CST	Binomial regression
Proportion of eyes with CST less than gender-specific spectral domain equivalent of 250 µm on Zeiss Stratus and at least a 10% CST decrease from baseline*	Binomial regression
Success proportion: 10% CST decrease from baseline	Binomial regression
Mean change in OCT retinal volume	ANCOVA

*No imputation of missing data because machine-specific, rather than time-domain equivalent, values are used

A plot that shows the mean change in OCT CST from baseline by treatment group over time will be constructed with observed data.

2.4 Exploratory Outcomes

2.4.1 Change in Visual Acuity Area Under the Curve

Change in visual acuity area under the curve (AUC) will be computed for each participant from baseline to 104 weeks using the imputed data sets. Analysis of covariance adjusting for baseline visual acuity and recent or planned DME treatment will be used to test for differences between the treatment groups.

Only common visits (baseline, 8, 52, and 104 weeks) will be used in the AUC analysis. AUC will be calculated for each participant according to the trapezoidal rule:

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

Where V_i is the change in visual acuity from baseline measured at the i^{th} visit, d is the number of days between visits i and $i+1$, and n is the number of common visits included in the analysis. For example, the 104-week outcome has $n = 4$ as the analysis will include visits at baseline, 8, 52, and 104 weeks. Note that the change in visual acuity from baseline is equal to 0 at baseline. For presentation, AUC will be divided by the number of days between baseline and the 104-week visit so that the value shown will have units of letters rather than letter·days. This statistic can then be interpreted as the average change in visual acuity over the time between baseline and the 104-week visit.

2.4.2 Diabetic Retinopathy

Several outcomes of interest related to diabetic retinopathy (DR) will be assessed at 52 and 104 weeks using observed data only (no imputation). All analyses will adjust for baseline DR severity and recent or planned DME treatment. The outcomes and the analysis techniques to be used are specified in Table 4, while definitions for DR improvement and worsening outcomes are given in Table 5.

Table 4. Analyses of Diabetic Retinopathy Severity.

Outcome	Analysis Technique
Proportion with improvement in DR severity graded on color fundus photographs	Binomial regression
Proportion of eyes with worsening of DR graded on color fundus photographs	Binomial regression
Time to worsening of DR on color fundus photographs over 2 years	Cox proportional hazards regression
Among eyes with PDR at randomization, development of vitreous hemorrhage or receipt of PRP, anti-VEGF for PDR, or vitrectomy for PDR	Cox proportional hazards regression

Table 5. Definitions for Improvement and Worsening of Diabetic Retinopathy on Photos

Baseline		Worsening (if FU ≥)	Improvement (if FU ≤)
NPDR	10/12	35	*Exclude*
	14/15/20	43	*Exclude*
	35	47	10/12
	43	53	14/15/20
	47	60	35
	53	60	43
PDR	60	65	*Exclude*
	61	71	**53 or 60***
	65	75	**53 or 60***
	71	81	61
	75	81	65
	81	*Exclude*	71
	85	*Exclude*	75

FU, follow up

*If an eye had PRP prior to baseline (on clinical exam), changing from ≥ 61 to 60 will be defined as improvement. If an eye did not have PRP prior to baseline, changing from 61/65 to 53 will be defined as improvement.

For the time-to-worsening analysis, participants who are lost to follow up will be considered censored on the day of their last visit. Participants that do not experience worsening of DR and complete the 104-week visit will be considered censored on the day of that visit. Hazard ratios will be presented along with the cumulative probability of worsening within each group to aid in interpretation.

A Kaplan-Meier curve showing time-to-worsening of DR by treatment group will be constructed along with the number of participants at risk at baseline and 52 weeks, and the number of events through 52 and 104 weeks.

2.4.3 Additional Exploratory Outcomes

The proportion of eyes with focal leakage patterns on fluorescein angiography randomized to prompt focal/grid photocoagulation with deferred intravitreous anti-VEGF that did not require subsequent anti-VEGF treatment will be tabulated.

3.0 Eyes in the Observational Phase

Eyes in the observational phase will be followed until one of the following occurs:

1. The eye is randomized
2. The eye receives non-topical DME treatment as part of usual care
3. The participant reaches 2 years (104 weeks) from enrollment

The primary objective of the observational phase is to collect data on the natural history of eyes that present with CI-DME and good vision that do not enroll in the randomized trial initially. Therefore, the proportion and 95% Wilson (Score) confidence interval (Newcombe 1998) of eyes that meet the following endpoints will be determined:

- Never need treatment
- Receive non-topical DME treatment
- Are randomized into Protocol V

Eyes completing 104 weeks of follow up and eyes lost to follow up in the observational phase that do not receive non-topical DME treatment or enter the randomized trial will be counted as never needing treatment.

3.1 Observational Phase Exploratory Analyses

In addition, data from the observational phase will be used in exploratory analyses to identify subgroups not needing DME treatment, to explore outcomes in eyes never needing treatment, and to explore outcomes in eyes randomized immediately vs. after enrolling in the observational phase. The specifics of these analyses will be developed later.

4.0 Economic Analysis

The purpose of the economic analysis is to compare the treatment groups with respect to cost and cost-effectiveness. The analysis plan is briefly described and will be detailed in a separate document.

Resource utilization data will be calculated using the number of clinic visits, along with the number and types of diagnostic and therapeutic ocular procedures performed on each group. To capture patient resource utilization, cost data for all diagnostic and therapeutic procedures performed will be tabulated to obtain a total cost for eye care services over 2 years of follow up. To capture the health-related quality-of-life associated with receipt of the three interventions over the course of the trial, two methods will be used. The first method will be to convert the visual acuities from the better-seeing eye over the two years of the trial into Quality-Adjusted Life-Years (QALYs) using the methods of Brown et al (2003). This method has been used widely in prior cost-effectiveness analyses of ophthalmologic interventions. The second method will use the best-corrected visual acuities from the treated eye, regardless of whether it is the better or worse-seeing eye. Resource use, costs and QALYS will be aggregated over the two years of the trial. The incremental cost-effectiveness ratio (ICER) will calculated by taking the incremental cost of prompt anti-VEGF over prompt focal/grid photocoagulation with deferred intravitreous anti-VEGF or observation with deferred anti-VEGF and dividing them by the incremental QALYs of prompt anti-VEGF over prompt focal/grid photocoagulation with deferred intravitreous anti-VEGF or observation with deferred anti-VEGF. A probabilistic sensitivity analysis will be conducted to better characterize overall uncertainty in the results.

5.0 Safety Analysis

Adverse events will be categorized as study eye, fellow eye, and systemic. All randomized eyes will be included in the safety analyses and analyzed according to treatment group assignment at randomization. An additional tabulation will be made for adverse events possibly related to study treatment. For the 104-week primary analysis, any adverse event that occurred at least once prior to the 104-week visit (or 728 days if the participant did not complete the 104-week visit) will be reported. If the overall test has $P \leq .05$, then pairwise comparisons between groups will be performed with no further adjustment for multiple comparisons. For all analyses, the hypothesis test of no difference between treatment groups will be conducted.

Due to the different visit schedules among the treatment groups, the ratio of adverse events and number of visits will be provided in addition to the number of eyes with an adverse event and the total number of adverse events for each treatment group. This will attempt to account for a potential disproportion of reported adverse events observed in the prompt anti-VEGF treatment group because of having more visits.

5.1 Ocular Adverse Events (Injection and Drug-Related)

Ocular adverse events will be tabulated separately for the three treatment groups. The frequency of the event occurring at least once per study eye will be calculated. Eye-level outcomes will be compared between treatment groups using Fisher's exact test.

The following adverse events will be assessed:

- Endophthalmitis
- Any retinal detachment (rhegmatogenous, tractional, combined rhegmatogenous and tractional, or not otherwise specified)
    - Rhegmatogenous retinal detachment (tabulated without formal analysis)
    - Tractional retinal detachment (tabulated without formal analysis)
- Retinal tear
- Cataract
- Cataract surgery
- Vitreous hemorrhage
- Ocular inflammation
- Intraocular pressure (IOP) elevation (any of the following)
    - Increase of IOP $\geq$ 10 mmHg from baseline (at a follow-up visit)
    - IOP $\geq$ 30 mmHg (at a follow-up visit)
    - Initiation of glaucoma medications
    - Glaucoma procedure
- Neovascularization of the iris or neovascular glaucoma

5.2 Systemic Adverse Events

Systemic adverse events will tabulated separately for the three treatment groups. The frequency of the event occurring at least once per participant will be calculated. Rates of systemic adverse events will be compared using Fisher's exact test.

Primary:

- Death
- Serious adverse event (at least one)
- Hospitalizations (at least one)
- Cardiovascular and cerebrovascular events according to the Antiplatelet Trialists' Collaboration (excerpted from BMJ Jan 8, 1994):
    - Non-fatal myocardial infarction
    - Non-fatal 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 (non-fatal myocardial infarction, non-fatal 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. Non-fatal myocardial infarction and non-fatal stroke require that the patient is alive at the end of the study. If not, then only the death is counted.

Secondary (for tabulation without formal statistical comparison):

- Hypertension
- Frequency of at least one event per participant in each Medical Dictionary for Regulatory Activities (MedDRA) system organ class

6.0 Additional Tabulations and Analyses

The following will be tabulated according to treatment group:

- Baseline demographic and clinical characteristics
- Visit completion rate for each visit
- Protocol deviations

The following treatment-related quantities will be tabulated by treatment group for patients that complete the 52- and 104-week visits:

- Number of injections
- Number of focal/grid photocoagulation sessions performed
- Number of visit (scheduled or unscheduled)

The proportion of eyes needing anti-VEGF treatment (deferred anti-VEGF groups only) also will be tabulated. All participants will be included in the calculation of the proportion, regardless of loss to follow up or visit completion.

7.0 Interim Monitoring Plan

A formal plan for interim data monitoring was established in consultation with the Data and Safety Monitoring Committee. The details are provided in the following document:

F:\user\aDRCRN\Protocols\Protocol V - Anti-VEGF vs Laser for CI DME with Excellent VA\Statistics\Interim analysis\Protocol V Interim Monitoring Plan v1.0 10-08-14.docx

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 6.

Table 6. Analysis Windows for Common Visits

Visit (Protocol Window)	Target	Analysis Window	
8 weeks ± 1 or 2 weeks*	56 days	28 – 84 days	(8 ± 4 weeks)
52 weeks ± 2 weeks	364 days	308 – 420 days	(52 ± 8 weeks)
104 weeks ± 4 weeks	728 days	644 – 812 days	(104 ± 12 weeks)

*Within 2 weeks for deferred anti-VEGF arms and within 1 week for prompt anti-VEGF arm

8.3 Missing Data

The strategy for handling missing data generally is included with the description of each 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 CST outcomes, these variables will be truncated to ± 3 standard deviations, with the standard deviation based upon 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. Change in visual acuity AUC will be calculated from imputed, truncated data, with 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, covariates may be categorized or excluded, and a non-parametric approach, robust method, or transformation may be considered.

9.0 Rationale for Key Changes

9.1 Version 1.0 to 2.0

The following changes were made in version 2.0. Changes were made prior to the primary analysis but after interim analysis. However, changes were not in response to the results of the analysis.

- Section 2.1
    - Removed the provision for using Poisson regression with an identity link if binomial regression with an identity link fails to converge. Poisson regression with the canonical log link will be used in this situation because it is more stable. In addition, Poisson regression with identity link was not pre-specified in the protocol.
    - Removed allocation of .001 alpha for DMSC review. The primary analysis will now be conducted with alpha of .05 rather than .049. The DSMC was supportive of this change, as such an allocation was considered arbitrary.
- Sections 2.1.1 to 2.1.5
    - Added provision that the sensitivity analysis with exact logistic regression will only be performed if there are fewer than 5 events in any of the treatment groups. The purpose of this analysis is to validate the results of the primary analysis if event rates are much lower than expected. If there are at least 5 events in each group, then the primary analysis is expected to be reliable.
    - Added provision that the per-protocol analysis will be performed only if more than 10% of randomized eyes would be excluded by the per protocol criteria. It is thought that any less than this is unlikely to have an effect on the results.
    - Added clarification that within-subgroup estimates of treatment effect will only be presented if the *P* value for the interaction term is $\leq .05$.
    - Increased the minimum number of eyes per treatment group per subgroup that will be required for a subgroup analysis to be performed from 10 to 20. Having fewer than 20 eyes in a group could lead to imprecise and unreliable estimates.
    - Reclassified leakage pattern on fluorescein angiogram as an exploratory subgroup analysis because subgroup categories are still being defined.
    - Added presence of circinate ring as an exploratory subgroup analysis because it is of interest to investigators.

- Added duration of diabetes, duration of DME, and lens status as exploratory subgroup analyses because they are listed in the protocol.

- Changed the minimum number of eyes per treatment group per center required for analysis of center effects from 30 to 20 to harmonize with the minimum number required for subgroup analyses.

- Sections 2.2 to 2.4

  - Clarified that the proportion of eyes with CST below gender- and machine-specific cutoffs will be analyzed without imputation of missing data because this outcome is calculated based on machine-specific values rather than the time-domain converted values that will be imputed for other analyses.

  - Added retinal volume as a secondary outcome for consistency with other protocols.

  - Changed analysis method from binomial regression to cox proportional hazards regression for development of a PDR event or receipt of treatment to treat a PDR event. Since the component outcomes can occur at any time, the time-to-event approach will increase statistical precision.

- Section 3.0

  - Replaced exact mid-*P* confidence intervals with Wilson (Score) confidence intervals for analysis of observational phase data because the exact intervals are unnecessarily conservative (Newcombe 1998). In addition, the Wilson interval has a simple formula for calculation.

- Section 5.0

  - Changed alpha level of safety analyses from .01 to .05. The DSMC was supportive of this change as safety analyses in this trial should be viewed as exploratory.

  - Added cataract as a safety outcome because it was pre-specified in the protocol.

  - Added neovascular glaucoma to form a composite safety outcome with neovascularization of the iris as these outcomes are on the same disease pathway (neovascularization of the iris is a precursor to neovascular glaucoma). In addition, event rates for both outcomes are expected to be low in this population.

  - Combined death of unknown cause with death attributed to cardiac, cerebral, hemorrhagic, embolic, or other vascular cause into one outcome for consistency with other protocols.

- Added composite APTC safety outcome for consistency with other protocols.

- Section 8.4
  - Modified the handling of outliers in the AUC analysis for consistency with other protocols.

References

Brown NM, Brown GC, Sharma S, Landy J. Health care economic analyses and value-based medicine. *Surv Ophthalmol*. 2003;48(2):204-23.

Hochberg Y. A Sharper Bonferroni Procedure for Multiple Significance Testing. *Biometrika*. 1988;75:800–803.

Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med.* 1998;17:857-872.

Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol*. 2005;162(3):199-200.

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*. 2016. DOI: https://doi.org/10.1177/0962280216683570.