Diabetic Retinopathy Clinical Research Network

Intravitreous Anti-VEGF vs. Prompt Vitrectomy for Vitreous Hemorrhage from Proliferative Diabetic Retinopathy (Protocol AB)

Statistical Analysis Plan

VERSION NUMBER	AUTHOR	APPROVER	EFFECTIVE DATE	REVISION DESCRIPTION
1.0	Wesley Beaulieu	Michele Melia	17 October 2017	Initial version for Protocol version 1.0.
1.1	Wesley Beaulieu	Michele Melia	12 February 2019	Revisions for consistency across DRCR.net SAPs following DSMC review. Still applies to Protocol version 1.0. Changes made prior to initial data analysis.
1.2	Wesley Beaulieu	Maureen Maguire	27 April 2020	Updated significance threshold for primary outcome from .05 to .049 for consistency with protocol. Added post hoc visual acuity subgroup analysis. Changed measure of treatment group effect for binary outcomes to risk difference instead of odds ratio. Other minor changes and clarifications. Changes were made after the initial analysis. Still applies to Protocol version 1.0.

SIGNATURES	
AUTHOR	Wesley Beaulieu I am digitally signing this document 2020-04-27 16:05-04:00
APPROVER	Maureen G. Maguire Digitally signed by Maureen G. Maguire Date: 2020.04.27 16:27:07 -04'00'

1.0 Introduction

This document outlines the statistical analysis plan for the Diabetic Retinopathy Clinical Research Network (DRCR.net) Protocol AB comparing prompt vitrectomy with panretinal photocoagulation (PRP) versus an intravitreous anti-vascular endothelial growth factor (VEGF) regimen for treatment of vitreous hemorrhage from proliferative diabetic retinopathy (PDR) for which intervention is deemed necessary. The anti-VEGF agent used in this trial is aflibercept (Eylea®, Regeneron Pharmaceuticals, Tarrytown, New York, USA).

The primary objective of the protocol is to determine if there is a difference between the two treatment groups in visual acuity area under the curve (AUC) over 24 weeks. Participants will have outcome visits at 4, 12, 24, 36, 52, 68, 84, and 104 weeks. The long-term outcomes (e.g., 2-year visual acuity AUC, other secondary outcomes, safety) are considered equally important as the primary outcome. Therefore, results from this trial will not be published until the 104-week follow-up has closed.

Study eyes will be assigned randomly to the two treatment groups in a 1:1 ratio stratified by site. Participants may have only one study eye enrolled in the randomized trial.

2.0 Efficacy Analysis Plan

2.1 Primary Outcome Analysis

The primary analysis will consist of a treatment group comparison of mean visual acuity AUC from baseline to 24 weeks adjusting for baseline visual acuity and phakic status using a general linear model with robust variance estimation. The primary analysis is an intention-to-treat analysis. All randomized eyes will be included in the primary analysis according to treatment group assignment at randomization. Only baseline and outcome visits (4, 12, 24, 36, 52, 68, 84, and 104 weeks), which are common to both groups, will be used to calculate AUC through the appropriate visit for the primary and secondary outcomes (e.g., 24 weeks for the primary outcome). AUC will be calculated for each participant by the trapezoidal rule using the following formula:

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

Where V_i is the visual acuity measured at the i^{th} visit, d is the number of days between visits i and $i+1$, and n is the number of outcome visits included in the analysis. For example, the primary outcome has $n = 4$ as the analysis will include visits at baseline, 4, 12, and 24 weeks. 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., 168 days for the primary outcome at 24 weeks). This statistic can then be interpreted as the average visual acuity over the time between baseline and the n^{th} visit. If the P value for the test of the treatment effect is less than or equal to .049, then it will be concluded that there is a significant difference in visual acuity AUC between the two groups.

Markov chain Monte Carlo (MCMC) multiple imputation will be used to impute missing data. The imputation model will include the visual acuities measured at baseline and at all outcome visits (4, 12, 24, 36, 52, 68, 84, and 104 weeks) along with treatment group and phakic status. If the data are judged to be substantially non-normal, then predictive mean matching with a match set of $k = 5$ may be used for imputation instead of MCMC. Due to the nature of vitreous hemorrhage, there is a higher than usual chance of the data following a non-normal distribution. To limit the influence of extreme data points, visual acuity will be truncated to ± 3 standard deviations based on observed 24-week data (see Section 7.4) prior to the AUC calculation but after imputation. Imputed data implying a nonbiologic value for visual acuity (< 0 or > 100) will be truncated at the closest biologic value (i.e., 0 or 100).

A plot showing the mean level of visual acuity by 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 numbers from statistical models (e.g., treatment group differences, confidence intervals, and *P* values) will be based on imputed data.

2.1.1 Sensitivity Analyses

A sensitivity analysis including observed data from participants completing the 24-week visit (no imputation of missing data) will be conducted (i.e., complete-case analysis). 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 differences. An additional sensitivity analysis using observed visual acuity data without truncation or imputation of missing data will also be performed.

A sensitivity analyses using nonparametric methods will be conducted using observed data from participants completing the 24-week visit. The primary outcome will be converted to van der Waerden (Normal) scores and analyzed similarly as above with a general linear model adjusting for baseline visual acuity and phakic status.

Multiple imputation assumes that data are missing at random (MAR). In the present study, this would mean that whether follow-up visual acuity data are missing or observed may be a function of observed baseline characteristics included in the imputation model (e.g., baseline visual acuity, follow-up visual acuity, treatment group, or phakic status), but not a function of the unobserved follow-up visual acuities that are being imputed. This assumption cannot be tested since these unobserved 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 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 conclusion of the primary hypothesis test changes. That is, if one group is found to be superior, the tipping point will identify the shift parameter necessary to nullify the result. Conversely, if the null hypothesis is not rejected, two tipping points will be identified – one that would make aflibercept superior and one that would make vitrectomy

superior. In either case, the tipping point(s) will be evaluated to determine plausibility. If not plausible, 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 for this analysis.

2.1.2 Per-Protocol Analysis

A per-protocol analysis will be conducted to estimate the treatment effect for each treatment among those who complied with the treatment. This analysis will include observed data (no imputation) from all randomized eyes that complete the 4-, 12-, and 24-week visits except those in the prompt vitrectomy group that never receive vitrectomy and eyes in the aflibercept group that missed at least one injection before the 24-week visit. 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 the factors that may have contributed to the differences. 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 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
- Mean arterial blood pressure
- Duration of vitreous hemorrhage
- HbA1c
- Prior PRP
- Prior treatment for diabetic macular edema (DME)

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.

2.1.4 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 will be biased towards the null hypothesis of no interaction (Sullivan et al., 2016). It is recognized that

analyzing only observed data may be biased, but unlike the imputed analysis, it is not automatically biased in the presence of interaction.

A significant 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 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, finding a significant treatment effect in a subgroup will be regarded as likely due to chance.

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

- History of PRP: yes vs. no
- Phakic status: phakic vs. pseudophakic
- Age: continuous and < 60 vs ≥ 60 years

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**
History of PRP	Eyes with prior PRP may have less active disease, which could be impacted differently by treatment. Alternatively, eyes with prior PRP that subsequently have a vitreous hemorrhage could have more severe underlying disease, which could be impacted differently by treatment.
Phakic status	Phakic eyes may have more cataract progression in the vitrectomy group than in the aflibercept group, which would reduce visual acuity.
Age	Eyes from older individuals may have a longer recovery time from surgery and be at a higher risk of cataract, which could be exacerbated in the vitrectomy group.

The following subgroup factors 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.

- Prior DME treatment: yes vs. no
- Prior focal/grid laser for DME: yes vs. no
- Prior anti-VEGF for DME: yes vs. no
- HbA1c: continuous and $< 7.5\%$ vs. $\geq 7.5\%$
- Duration of vitreous hemorrhage: below vs. at or above median
- Sex: female vs. male
- Race/Ethnicity: White vs. Black/African American vs. Hispanic (exclude all other groups due to anticipated small sample size) and White vs. non-White
- Baseline visual acuity: 20/32 to 20/800 (78 to 4 letters) vs. worse than 20/800 (< 3 letters)
    - The subgroup analysis by baseline visual acuity was added after reviewing the data (i.e., post hoc)

To increase the accuracy of estimates, subgroups will only be analyzed if there are at least 20 eyes in each treatment group for each subgroup. Cutoffs of continuous and ordinal outcomes may be modified to achieve a reasonable number of eyes in each group. Interaction *P* values will be calculated using the continuous and 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 = 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.

2.2 Secondary Outcome Analyses

2.2.1 Visual Acuity

Additional analyses of visual acuity will use the imputed data sets created for the primary outcome. These analyses are summarized in Table 2. The analysis models will include adjustments for baseline visual acuity and phakic status. Analyses will be conducted at 4, 12, 24, 52, and 104 weeks unless otherwise noted. Note that visual acuity AUC at 24 weeks is the primary outcome and not a secondary outcome.

For all outcomes analyzed with logistic regression, the percentage of eyes meeting the outcome at the visit will be reported for each treatment group. The treatment group comparison will be summarized with a risk difference and 95% confidence interval (estimated using the delta method, see Localio et al. 2007) and a *P* value. For binary change in visual acuity outcomes, only eyes at risk of the outcome will be included in the analysis. For example, because the minimum visual acuity letter score is 0 letters, only eyes with baseline visual acuity ≥ 15 letters will be included in the analysis ≥ 15-letter loss. All eyes will be included in the analyses of binary visual acuity letter score (e.g., ≥ 69 letters) since visual acuity may increase or decrease following randomization, which makes all eyes at risk for the outcome.

Table 2. Visual Acuity Secondary Outcome Analyses.

Outcome	Analysis Technique
Visual acuity letter score AUC (104 weeks only)	General linear model
Visual acuity letter score	General linear model
Success proportion: visual acuity ≥ 84 letters (~20/20)	Tabulation only
Success proportion: visual acuity ≥ 74 letters (~20/32)	Logistic regression
Success proportion: visual acuity ≥ 69 letters (~20/40)	Tabulation only
Failure proportion: visual acuity ≤ 38 letters (~20/200)	Logistic regression
Failure proportion: visual acuity ≤ 8 letters (~20/800)	Tabulation only
Success proportion: visual acuity gain ≥ 15 letters *	Logistic regression
Success proportion: visual acuity gain ≥ 30 letters *	Tabulation only
Failure proportion: visual acuity loss ≥ 15 letters *	Logistic regression
Failure proportion: visual acuity loss ≥ 30 letters *	Tabulation only

* These outcomes are considered exploratory because they were added after the protocol was finalized.

2.2.2 PDR and DME

Additional secondary outcomes will be evaluated at 24, 52, and 104 weeks. Analyses will use observed data (no imputation) and adjust for phakic status.

Table 3. Additional Secondary Outcomes.

Outcome	Analysis Technique
Percentage of eyes with recurrent vitreous hemorrhage on clinical exam at any time during follow-up*	Logistic regression
Percentage of eyes with retinal neovascularization on clinical exam (neovascularization of the disc or elsewhere)	Logistic regression
OCT central subfield thickness†	General linear model

*A subgroup analysis similar to what is described in section 2.1.4 will be conducted for this outcome by whether the participant was on anticoagulants or antiplatelet medication at baseline. This outcome will be evaluated through 104 weeks only.

†No adjustment for baseline central subfield thickness because vitreous hemorrhage will prevent obtaining usable OCT scans for many participants.

A plot of mean OCT central subfield thickness over time by group will be constructed using observed data.

2.2.3 Workplace Productivity and Activity Impairment Questionnaire

Outcomes from the Workplace Productivity and Activity Impairment Questionnaire (WPAIQ) will be compared between treatment groups at 4, 12, 24, 52 and 104 weeks unless otherwise indicated in the bulleted list below. Analyses will be conducted with a general linear model that is adjusted for the baseline level of the score being analyzed and phakic status. Only participants completing the corresponding visit will be included in the analysis. There will be no imputation of missing data. The following outcomes will be evaluated:

- Change in Absenteeism score (tabulated without statistical comparison)
- Change in Presenteeism score (tabulated without statistical comparison)
- Change in Work Productivity Loss score
- Change in Work Productivity Loss score AUC (24 and 104 weeks only)
- Change in Activity Impairment score
- Change in Activity Impairment score AUC (24 and 104 weeks only)

WPAIQ AUC scores will be calculated similarly to visual acuity AUC (described in Section 2.1).

The Work Productivity Loss score (a composite of the Absenteeism and Presenteeism scores) will be considered the scale of primary interest from the WPAIQ. It is anticipated that the vitrectomy group will be superior to the anti-VEGF group due to a faster recovery of visual acuity. Plots showing the mean levels of the Work Productivity Loss and Activity Impairment scores by treatment group over time will be constructed.

2.3 Exploratory Outcomes

The following outcomes were not outlined in the protocol. However, they are considered of interest. These analyses will adjust for baseline phakic status and use observed data only.

Table 4. Exploratory Outcomes.

Outcome	Analysis Technique
Presence of DME[*]	Logistic regression
Time to initiation of DME treatment through 104 weeks[†]	Cox Proportional Hazards Regression

[*]Defined as OCT central subfield thickness ≥ 305 µm for women and ≥ 320 µm for men (Heidelberg Spectralis) or ≥ 290 for women and ≥ 305 for men (Zeiss Cirrus). Evaluated at 24, 52, and 104 weeks for eyes completing the visit.
[†]Defined as anti-VEGF injection, corticosteroid injection, administration of focal/grid laser, or vitrectomy for DME. Includes all randomized eyes. The hazard ratio with 95% confidence interval and *P* value will be reported for the full 104-week follow-up. A Kaplan-Meier plot will be constructed and the cumulative probabilities with 95% confidence intervals will be presented for each treatment group at 104 weeks.

The ability to obtain an OCT scan may be related to the treatment if the hemorrhage clears more quickly in one arm, potentially introducing bias into the treatment comparison for these two outcomes. Therefore, the proportion of eyes in which an OCT scan cannot be obtained will be tabulated for each group at 24, 52, and 104 weeks.

2.3.1 Development of DME

Development of DME in eyes with OCT assessment at baseline with be evaluated as an exploratory outcome over the full 2-year follow-up. Having DME on OCT at baseline is an exclusion criterion for randomization. It is recognized that OCT will be unavailable for many eyes due to dense vitreous hemorrhage at baseline; however, since OCT is obtained prior to randomization, the treatment comparison is still valid. The analysis will adjust for baseline OCT central subfield thickness and phakic status. Therefore, only eyes with available baseline OCT data will be included. Cox proportional hazards regression will be used to estimate the hazard ratio and associated 95% confidence interval for developing DME with vitrectomy versus aflibercept. Data from eyes not developing DME within the 2-year follow-up period will be considered censored on the day of their final completed visit, without regard for interim missed visits. To increase the accuracy of estimates, this analysis will be conducted only if at least 20 eyes in each treatment group meet the outcome criteria.

A Kaplan-Meier curve showing time-to-development of DME by treatment group will be constructed. Within each treatment group, the Kaplan-Meier estimate of the survival function and 95% confidence interval at the final available time point will be calculated.

3.0 Outcomes within Treatment Groups

Within each treatment group, the following outcomes will be tabulated at 24, 52, and 104 weeks (for 24-, 52-, and 104-week completers, respectively), both cumulatively from baseline, and for the intervening periods between those visits. There will be no formal statistical comparisons.

- Proportion of eyes undergoing vitrectomy (initial vitrectomy in aflibercept group and repeat vitrectomy in vitrectomy group)
- Distribution and mean (standard deviation) number of intravitreous aflibercept injections performed
- Proportion of eyes receiving PRP during follow-up

For the vitrectomy group, the following components of the surgical treatment will be characterized with data tabulations and descriptive statistics:

- Surgery time in minutes
- Vitrectomy system gauge
- Intraoperative use of anti-VEGF or steroid
    - Drugs used

- Administration of PRP
    - For those undergoing PRP, also tabulate the average power (mW), exposure time (seconds), total number of spots, estimated number of quadrants treated, estimated total number of burns that had a visible effect on the retina, and if the PRP is considered "complete" as defined by the protocol
- Surgical complications (e.g., choroidal detachment, rhegmatogenous retinal detachment, retinal dialysis, retinal tear, etc.)
- Use of staining agents to visualize epiretinal proliferation or internal limiting membrane
    - Agent used
- Placement of gas in the eye (C_3F_8, SF_6, or filtered air)
- Additional procedure(s) performed during vitrectomy surgery (e.g., cataract extraction, cryopexy, fluid-gas exchange, epiretinal membrane peeling, internal limiting membrane peel, intraocular lens implantation, laser retinopexy, removal of fibrous proliferation, scleral buckle, etc.)

4.0 Economic Analysis

The purpose of the economic analysis is to compare the treatment groups with respect to cost, cost effectiveness, and workplace productivity loss. 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 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 two 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 intravitreous aflibercept over vitrectomy and dividing it by the incremental QALYs of aflibercept over vitrectomy. 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, or systemic. A full listing of adverse events will be tabulated by treatment group for systemic, study-eye ocular, and fellow-eye ocular adverse events. 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 the treatment group they were assigned to at randomization. For the 104-week primary safety 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.

Due to the different visit schedules among the treatment groups, the mean ratio of adverse events divided by the 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 aflibercept group because of having more visits.

5.1 Study Eye Ocular Adverse Events

The proportion of eyes experiencing each event listed below will be reported for study eyes by treatment group. Proportions will be compared between treatment groups using Barnard's unconditional exact test, considering the number of eyes randomized to each group as fixed.

The following ocular adverse events will be assessed:

- Endophthalmitis
- Any retinal detachment (rhegmatogenous, traction, combined rhegmatogenous and traction, or not otherwise specified)
    - Rhegmatogenous retinal detachment (tabulated without statistical comparison)
    - Traction retinal detachment (tabulated without statistical comparison)
- Retinal tear
- Cataract extraction in eyes phakic at baseline or visually significant cataract on clinical exam as defined on the case report form
    - Cataract extraction in eyes phakic at baseline (tabulated without statistical comparison)
    - Visually significant cataract on clinical exam as defined on the case report form (tabulated without statistical comparison)
- Ocular inflammation
- Intraocular pressure (IOP) elevation (any of the following)
    - Increase in IOP $\geq$ 10 mmHg from baseline (at a follow-up visit)
    - IOP $\geq$ 30 mmHg (at a follow-up visit)
    - Initiation of medication to lower IOP that was not in use at baseline
    - Glaucoma procedure
- Neovascular glaucoma

- Neovascularization of the iris

5.2 Systemic Adverse Events

The proportion of randomized participants experiencing each event listed below will be reported by treatment group. Proportions will be reported and compared using Barnard's unconditional exact test, considering the number of eyes randomized to each group as fixed. The following serious systemic adverse events will be assessed:

- Primary:
  - Death
  - Serious adverse event (at least one)
  - Hospitalization (at least one)
  - 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 is alive at the end of the study. If not, then only the death is counted.

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

6.0 Additional Tabulations

The following will be tabulated according to treatment group:

- Baseline demographic and clinical characteristics
- Visit completion rate for each outcome visit (4, 12, 24, 36, 52, 68, 84, and 104 weeks)
- Treatment completion

7.0 General Principles for Analysis

7.1 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.

7.2 Visit Windows for Analysis

For common visits, the analysis windows will be defined according to Table 5. For visits falling in more than 1 window, priority will be given to the 104-, 52-, and 24-week visits. Otherwise, the visit will be assigned to the earlier window (e.g., a visit on day 42 would be assigned as the 4-week visit).

Table 5. Analysis Windows for Outcome Visits

Visit (Protocol Window)	Target	Analysis Window	
4 (±1) weeks	28 days	14 – 42 days	(4 ± 2 weeks)
12 (±4) weeks	84 days	42 – 126 days	(12 ± 6 weeks)
24 (±4) weeks	168 days	126 – 210 days	(24 ± 6 weeks)
36 (±4) weeks	252 days	210 – 308 days	(36 +8 weeks)
52 (±4) weeks	364 days	308 – 420 days	(52 ± 8 weeks)
68 (±4) weeks	476 days	420 – 532 days	(68 ± 8 weeks)
84 (±4) weeks	588 days	532 – 644 days	(84 ± 8 weeks)
104 (±4) weeks	728 days	644 – 812 days	(104 ± 12 weeks)

7.3 Missing Data

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

7.4 Outliers

To help ensure that statistical outliers do not have undue impact on analyses of continuous visual acuity and OCT central subfield thickness outcomes (including the primary outcome), outcomes will be truncated to ± 3 standard deviations based on the mean and standard deviation at 24 weeks for 24-week completers, irrespective of treatment group. Visual acuity letter score, change in visual acuity from baseline, and OCT central subfield thickness will be truncated. 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.

7.5 Model Assumptions

All model assumptions, including linearity, normality of residuals, and heteroscedasticity, will be verified. If model assumptions are not reasonably satisfied, then covariates may be categorized or excluded, and a nonparametric approach, robust estimation method, or transformation may be considered. The possibility of using a transformation, robust method, or nonparametric approach is higher than usual in this study due to the nature of the cohort as vitreous hemorrhage severity can vary greatly, potentially leading to atypical (non-Gaussian) distributions of outcome data.

7.6 Type I Error Rate

There is no formal adjustment for multiplicity to compensate for the large number of outcomes being compared. All comparisons are conducted at alpha level 0.05 unless otherwise noted (e.g., the primary outcome is conducted at alpha level 0.049). In some cases, select secondary outcomes have been designated as being of "primary interest" along with a hypothesized direction of effect to place more weight on these outcomes over others.

References

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

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):874-82.

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.