

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

761355Orig1s000

STATISTICAL REVIEW(S)



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #: BLA 761355

Drug Name: Aflibercept 8 mg intravitreal (IVT) injection

Indication(s): For the Treatment of:

- Neovascular (wet) Age-Related Macular Degeneration (nAMD)
- Diabetic Macular Edema (DME) and
- Diabetic Retinopathy (DR)

Applicant: Regeneron Pharmaceuticals, Inc.

Date(s): Stamp Date: December 27, 2022
PDUFA Date: June 27, 2023
Review Date: May 26, 2023

Review Priority: Priority

Biometrics Division: IV

Statistical Reviewer: Juan C. Vivar, Ph.D.

Concurring Reviewers: Solomon Chefo, Ph.D., Team Leader

Medical Division: Division of Ophthalmology

Clinical Team: Wiley Chambers, MD, Medical Officer

Project Manager: Michael Puglisi

Keywords: Clinical Studies, MMRM, Diabetic Macular Edema, Neovascular Age-related Macular Degeneration, Diabetic Retinopathy, Best Corrected Visual Acuity

Table of Contents

1. EXECUTIVE SUMMARY	6
2. INTRODUCTION	10
2.1. OVERVIEW.....	10
2.1.1. <i>Specific Studies Reviewed</i>	11
2.2. DATA SOURCES	12
3. STATISTICAL EVALUATION	12
3.1. DATA SOURCES	12
3.2. EVALUATION OF EFFICACY	12
3.2.1. <i>PULSAR Study for nAMD Indication</i>	13
3.2.1.1. Study Design and Endpoints.....	13
3.2.1.2. Statistical Methodologies	15
3.2.1.3. Patient Disposition, Demographic and Baseline Characteristics.....	19
3.2.1.4. Results and Conclusions.....	22
3.2.1.5. Efficacy Conclusion	30
3.2.2. <i>PHOTON Study for DME/DR Indications</i>	31
3.2.2.1. Study Design and Endpoints.....	31
3.2.2.2. Statistical Methodologies	33
3.2.2.3. Patient Disposition, Demographic and Baseline Characteristics.....	39
3.2.2.4. Results and Conclusions.....	42
3.2.2.5. Efficacy Conclusion	54
3.3. EVALUATION OF SAFETY	55
3.3.1. <i>PULSAR Study</i>	55
3.3.2. <i>PHOTON Study</i>	57
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS.....	60
4.1. PULSAR STUDY FOR NAMD INDICATION.....	60
4.2. PHOTON STUDY FOR DME/DR INDICATIONS	61
5. SUMMARY AND CONCLUSIONS	63
5.1. STATISTICAL ISSUES	63
5.2. COLLECTIVE EVIDENCE.....	63
5.3. CONCLUSION AND RECOMMENDATION.....	64
5.4. LABELING RECOMMENDATION	65

LIST OF TABLES

Table 1: Adjusted Mean Change in BCVA from Baseline at Week 48 (FAS) (PULSAR)	7
Table 2: Adjusted Mean Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)	9
Table 3: Summary of Intercurrent Events Through Week 48 in PULSAR study	17
Table 4: Summary of Subject Disposition and Reasons for Study Discontinuation (All Randomized Subjects)	19
Table 5: Summary of Analysis Populations in PULSAR Study.....	20
Table 6: Demographic and Baseline Characteristics in PULSAR Study (Full Analysis Set)	20
Table 7: Number of Subjects with Observed BCVA Data by Visit (PULSAR)	22
Table 8: Summary of Mean Change in BCVA from Baseline at Week 48 (FAS) (PULSAR)	22
Table 9: Summary of Mean Change in BCVA from Baseline at Week 40, 44, 48, and averaged at Week 40/44/48 (FAS) (PULSAR).....	24
Table 10: Sensitivity and Supportive Analyses: Mean Change in BCVA from Baseline at Week 48 (PULSAR)	24
Table 11: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 16, LOCF (FAS)	25
Table 12: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 48 (FAS) (PULSAR)	26
Table 13: Proportion of Subjects Who Gained ≥ 15 , ≥ 10 , ≥ 5 and ≥ 0 Letters and Avoided Losing ≥ 5 , ≥ 10 and ≥ 15 Letters in BCVA from Baseline at Week 48 (FAS) (PULSAR).....	27
Table 14: Least Square Mean Change in CST from Baseline at Week 48, FAS (PULSAR)	29
Table 15: Summary of Intercurrent Events Through Week 48 in PHOTON study	35
Table 16: Summary of Subject Disposition and Reasons for Study Discontinuation (PHOTON) (All Randomized Subjects).....	39
Table 17: Summary of Analysis Populations (PHOTON)	39
Table 18: Summary of Demographic and Baseline Characteristics (FAS) (PHOTON)	40
Table 19: Number of Subjects with Observed BCVA Data by Visit (PHOTON)	42
Table 20: Adjusted Mean Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)	42
Table 21: Summary of Mean Change in BCVA from Baseline at Week 40, 44, 48, and averaged at Week 40/44/48 (FAS) (PHOTON)	44
Table 22: Sensitivity and Supportive Analyses: Mean Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)	44
Table 23: Number of Subjects with Observed DRSS Data from Baseline through Week 48.....	46
Table 24: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline to Week 48 (Full Analysis Set)	47
Table 25: Number of Subjects by Baseline DRSS category and by treatment (FAS) (PHOTON)	48
Table 26: Proportion of Subjects who Achieved ≥ 2 -Step Improvement in DRSS from Baseline to Week 48 by Baseline DRSS Category (FAS) (PHOTON).....	49
Table 27: Proportion of subjects Who Achieved ≥ 2 -Step Improvement from Baseline in the DRSS at Year 1.....	49

Table 28: Proportion of Subjects Who Gained ≥ 15 , ≥ 10 , ≥ 5 and ≥ 0 Letters and Avoided Losing ≥ 5 , ≥ 10 and ≥ 15 Letters in BCVA from Baseline at Week 48, FAS (PHOTON).....	50
Table 29: Proportion of Subjects with no Fluid at Foveal Center at Week 48 and Week 12, LOCF (FAS) (PHOTON)	52
Table 30: Least Square Mean Change in CST from Baseline at Week 48, FAS (PHOTON).....	53
Table 31: Summary of Study Treatment Exposure in the Study Eye through Week 48 (Safety Analysis Set) (PULSAR).....	55
Table 32: Overview of Safety Through Week 48 from PULSAR Study (Safety Analysis Set) ...	56
Table 33: Summary of Study Treatment Exposure in the Study Eye through Week 48 (Safety Analysis Set) (PHOTON).....	57
Table 34: Overview of Safety Through Week 48 from PULSAR Study (Safety Analysis Set) ...	58
Table 35: Adjusted Mean Change in BCVA from Baseline Through Week 48 by Treatment (Hypothetical Strategy - FAS) (PULSAR)	68
Table 36: Change from baseline in BCVA at Week 48, ANCOVA, Multiple Imputation with Tipping Point Analysis (FAS) (PULSAR)	68
Table 37: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 16. (Sensitivity and Supportive Analyses) (PULSAR).....	69
Table 38: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 48, Observed cases, FAS (PULSAR)	69
Table 39: Proportion of Subjects with no IRF and no SRF by Visit, LOCF (PULSAR).....	70
Table 40: BCVA-related Binary Efficacy Endpoints – Multiple Imputation (PULSAR).....	71
Table 41: Adjusted Mean Change in BCVA from Baseline Through Week 48 by Treatment (PHOTON)	72
Table 42: Change from baseline in BCVA at Week 48, ANCOVA, Multiple Imputation with Tipping Point Analysis (PHOTON)	72
Table 43: BCVA-related Binary Efficacy Endpoints – Multiple Imputation (PHOTON).....	73
Table 44: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline to Week 48 – Multiple Imputation (Full Analysis Set).....	74
Table 45: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline to Week 48 – Observed Cases (Full Analysis Set)	74
Table 46: Proportion of Subjects with no Fluid at Foveal Center at Week 48, OC (Full Analysis Set).....	75
Table 47: Proportion of Subjects with No Fluid in the Foveal Center by Visit, LOCF	75
Table 48: DRM Criteria in PULSAR and PHOTON Studies During Year 1	76

LIST OF FIGURES

Figure 1: Number of Subjects Enrolled by Region. Top panel: PULSAR study. Bottom panel: PHOTON study.	11
Figure 2: Dosing schedule for PULSAR study.....	14
Figure 3: Plot of Adjusted Mean Change in BCVA from Baseline Through Week 48	23
Figure 4: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (FAS) (PULSAR)	27

Figure 5: Proportion of Subjects Who Gained ≥ 15 Letters from Baseline through Week 48 based on LOCF (FAS) (PULSAR)	28
Figure 6: Least Square Mean Changes from Baseline in CST (in μm) by visit, MMRM, FAS (PULSAR)	29
Figure 7: Dosing schedule for PHOTON study.....	32
Figure 8: Graphical alpha-propagation approach (PHOTON)	37
Figure 9: Plot of Adjusted Mean change in BCVA from Baseline through Week 48 (FAS) (PHOTON)	43
Figure 10: Distribution of DRSS at Baseline in PHOTON Study (FAS).....	45
Figure 11: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline at each Visit (LOCF) (FAS).....	47
Figure 12: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)	50
Figure 13: Proportion of Subjects Who Gained ≥ 15 Letters from Baseline through Week 48 based on LOCF (FAS) (PHOTON)	51
Figure 14: Least Square Mean Changes from Baseline in CST (in μm) by visit, MMRM, FAS (PHOTON)	53
Figure 15: Adjusted Mean Change in BCVA from Baseline at Week 48 by Subgroup (Full Analysis Set) (PULSAR).....	60
Figure 16: Adjusted Mean Change in BCVA from Baseline at Week 48 by Subgroup	61
Figure 17: Proportion of Subjects Who Achieved ≥ 2 -Step Improvement in DRSS from Baseline at Week 48 by Subgroup (Full Analysis Set) (PHOTON)	62
Figure 18: Proportion of Subjects with no IRF and no SRF by Visit, LOCF (PULSAR).....	69
Figure 19: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (PULSAR).	70
Figure 20: Proportion of subjects who achieved ≥ 15 letters – Multiple Imputation (PULSAR)	71
Figure 21: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (PHOTON).	73
Figure 22: Cumulative Distribution of the Change in DRSS from Baseline at Week 48 (FAS) (PHOTON)	74
Figure 23: Proportion of Subjects with No Fluid in the Foveal Center by Visit, LOCF	75

1. EXECUTIVE SUMMARY

In this Biologics License Application (BLA), the Applicant seeks approval of a high dose (8 mg) of aflibercept intravitreal (IVT) injection administered every 12 weeks and every 16 weeks for the treatment of subjects with neovascular (wet) age-related macular degeneration (nAMD), diabetic macular edema (DME), and diabetic retinopathy (DR). Of note, aflibercept 2 mg IVT injection administered every 8 weeks after three initial monthly doses was approved for the treatment of nAMD and after five initial monthly doses was approved for the treatment of DME/DR.

nAMD Indication

In this BLA submission, efficacy and safety support for the nAMD indication was based on 48-week data from an ongoing pivotal Phase 3, 96-week, multi-center, randomized, double-masked, active-controlled noninferiority global study: PULSAR (Study 20968). The primary objective of PULSAR was to assess whether treatment with a high dose (HD) of aflibercept 8 mg at intervals of 12 or 16 weeks provide comparable efficacy benefit compared to aflibercept 2 mg dosed every 8 weeks.

In the PULSAR study, a total of 1011 treatment-naïve subjects at least 50 years of age, who met all the enrollment criteria were randomized in a 1:1:1 ratio and were to receive aflibercept 2 mg every 8 weeks after three initial monthly doses (hereafter referred to as 2q8), aflibercept 8 mg every 12 weeks after three initial monthly doses (hereafter referred to as HDq12), or aflibercept 8 mg every 16 weeks after three initial monthly doses (hereafter referred to as HDq16). Randomization was stratified by baseline best-corrected visual acuity (BCVA; <60 letters vs. ≥60 letters) and geographical region (Japan vs. Rest of the World).

Although the total study duration is 96 weeks, efficacy and safety support in this BLA was based on the first 48 weeks of data. The main efficacy evaluation was based on BCVA assessed every 4 weeks through Week 96 as measured by the number of letters read at a starting distance of 4 meters (range: 0-100 letters). The primary efficacy endpoint was the change in BCVA from baseline at Week 48. The primary efficacy analysis was an evaluation of noninferiority of HDq12 or HDq16 to 2q8 in the primary efficacy endpoint on the full analysis set (FAS), including all randomized subjects who received at least one dose of study medication (see more details on the data handling approaches in [Section 3.2.1.2](#)). The noninferiority margin was set at -4.0 letters.

The key secondary efficacy endpoint in the PULSAR study was the proportion of subjects without intraretinal fluid (IRF) and subretinal fluid (SRF) in central subfield at Week 16. The secondary efficacy analysis was an evaluation of superiority of HD of aflibercept (i.e., pooled HDq12 and HDq16) to 2q8 in the key secondary efficacy endpoint on the FAS. Superiority in the key secondary endpoint was tested if noninferiority of both HDq12 and HDq16 to 2q8 in the primary endpoint was established.

The treatment groups in PULSAR study were well balanced with respect to the demographic and baseline characteristics, in general (Table 6). During the first 48 weeks of treatment, about 7% of subjects discontinued from the study and the discontinuation rates were comparable across the treatment groups (Table 4).

In PULSAR study, both HDq12 and HDq16 treated subjects had a noninferior mean change in BCVA from baseline at Week 48 compared to 2q8 treated subjects (Table 1). As shown, the adjusted mean change in BCVA from baseline at Week 48 was +6.1 letters in the HDq12 group, +5.9 letters in the HDq16 group, and +7.0 letters in the 2q8 group with treatment differences of -1.0 (95% CI: -2.9 to 0.9) for HDq12 vs. 2q8 and -1.1 (95% CI: -3.0 to 0.7) for HDq16 vs. 2q8. Although noninferiority in the primary endpoint was established in the study, because of the frequent dosing, subjects that received aflibercept 2q8 displayed a numerical advantage in the BCVA gain from baseline through Week 48 compared to subjects that received the high dose of aflibercept (Figure 3).

Table 1: Adjusted Mean Change in BCVA from Baseline at Week 48 (FAS) (PULSAR)

Analysis Visit		2q8 (N = 336)	HDq12 (N = 335)	HDq16 (N = 338)
Baseline	Mean (SD)	59.0 (14.01)	59.8 (13.42)	60.1 (12.40)
Week 48	LS Mean (95% CI)	7.0 (5.6, 8.5)	6.1 (4.5, 7.6)	5.9 (4.5, 7.3)
	Diff (95% CI)	---	-1.0 (-2.9, 0.9)	-1.1 (-3.0, 0.7)

Note: Adjusted mean changes in each treatment group and treatment differences (95% CI estimates) were based on MMRM method using FAS (See Section 3.2.1.2 for details).

Sensitivity and supplementary analyses performed under different data handling strategies for missing data (including tipping point analysis) and intercurrent events were consistent with the primary efficacy results, leading to the same conclusion for a robust interpretation of the noninferiority findings (Table 10).

Additionally, PULSAR study results for the proportion of subjects who had no retinal fluid at Week 16 established superiority of the pooled HD groups compared to 2q8 group. For example, at Week 16, 63% of subjects in the pooled HD groups had no retinal fluid compared to 52% of subjects in the 2q8 group with a treatment difference of 12% (95% CI: 5%, 18%). Although superiority was established at Week 16, the efficacy results were not consistent throughout the study (Table 39 and Figure 18 in Appendix 1). This was likely because treatment groups had variable dosing intervals after Week 16 and an increasing number of subjects in each treatment group exhibited fluid dryness in the immediate visit after injection was received. Thus, the clinical benefit of HD of aflibercept in retinal fluid dryness compared to 2q8 is questionable. The reviewer defers to the clinical review team for the clinical meaningfulness of this endpoint.

In summary, based on the collective efficacy evidence from the adequate and well controlled trial of PULSAR study, the reviewer concludes that the application for the nAMD indication provided substantial evidence for comparable efficacy benefit (BCVA gain) of aflibercept 8 mg administered every 12 weeks and every 16 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after three initial monthly doses. However, the study did not establish consistent treatment benefit of the high doses of aflibercept in retinal fluid dryness compared to the active control.

DME/DR Indications

In this BLA submission, efficacy and safety support for the DME/DR indications was based on 48-week data from an ongoing pivotal Phase 3, 96-week, multi-center, randomized, double-masked, active-controlled noninferiority global study: PHOTON (Study VGFTe-HD-DME-1934). The primary objective of PHOTON was to assess whether treatment with high dose (HD) of aflibercept 8 mg at intervals of 12 or 16 weeks provided comparable efficacy benefit compared to aflibercept 2 mg dosed every 8 weeks.

In the PHOTON study, a total of 970 diabetic subjects at least 18 years of age, who met all the enrollment criteria were randomized in a 1:2:1 ratio and were to receive aflibercept 2 mg every 8 weeks after five initial monthly doses (2q8), aflibercept 8 mg every 12 weeks after three initial monthly doses (HDq12), or aflibercept 8 mg every 16 weeks after three initial monthly doses (HDq16). Randomization was stratified by baseline central retinal thickness (CRT; <400 μ m vs. $\geq$ 400 μ m), prior use of DME treatment (yes vs. no), and geographical region (Japan vs. Rest of the World).

Although the total study duration is 96 weeks, efficacy and safety support in this BLA was based on the first 48 weeks of data. The main efficacy evaluation was based on BCVA assessed every 4 weeks through Week 96 as measured by the number of letters read at a starting distance of 4 meters (range: 0-100 letters) and based on diabetic retinopathy severity as measured by the diabetic retinopathy severity score (DRSS).

The primary efficacy endpoint was the change in BCVA from baseline at Week 48. The primary efficacy analysis was an evaluation of noninferiority of HDq12 or HDq16 to 2q8 in the primary efficacy endpoint on the full analysis set (FAS), including all randomized subjects who received at least one dose of study medication (See more details in [Section 3.2.2.2](#)). The noninferiority margin was set at -4.0 letters.

The key secondary efficacy endpoint was the proportion of subjects with a $\geq$ 2-step improvement in DRSS at Week 48. This endpoint is intended to support the DR indication. The secondary efficacy analysis was an evaluation of noninferiority of each high dose group compared to the 2q8 group in the key secondary efficacy endpoint on the FAS. The noninferiority margin was set at -10%. Noninferiority in the key secondary endpoint was tested if noninferiority of both HDq12 and HDq16 to 2q8 in the primary endpoint was established.

The treatment groups in the PHOTON study were well balanced with respect to the demographic and baseline characteristics, in general ([Table 18](#)). During the first 48 weeks of treatment, about 7% of subjects discontinued from the study and discontinuation rates were comparable across the treatment groups ([Table 16](#)).

In PHOTON study, both HDq12 and HDq16 treated subjects had a noninferior mean change in BCVA from baseline at Week 48 compared to 2q8 treated subjects ([Table 2](#)). As shown, the adjusted mean change in BCVA from baseline at Week 48 was +8.1 letters in the HDq12 group, +7.2 letters in the HDq16 group, and +8.7 letters in the 2q8 group with treatment differences of -0.6 (95% CI: -2.3 to 1.1) for HDq12 vs. 2q8 and -1.4 (95% CI: -3.3 to 0.4) for HDq16 vs. 2q8.

Although noninferiority in the primary endpoint was established in the study, because of the relatively frequent dosing, subjects that received aflibercept 2q8 displayed a numerical advantage in the BCVA gain from baseline through Week 48 compared to subjects that received either high dose of aflibercept ([Figure 9](#)).

Table 2: Adjusted Mean Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)

Analysis Visit		2q8 (N = 167)	HDq12 (N = 328)	HDq16 (N = 163)
Baseline	Mean (SD)	61.5 (11.22)	63.6 (10.10)	61.4 (11.76)
Week 48	LS Mean (95% CI)	8.7 (7.2, 10.1)	8.1 (6.9, 9.3)	7.2 (5.8, 8.6)
	Diff (95% CI)	---	-0.6 (-2.3, 1.1)	-1.4 (-3.3, 0.4)

Note: Adjusted mean changes in each treatment group and treatment differences (95% CI estimates) were based on MMRM method using FAS (See [Section 3.2.2.2](#) for details).

Sensitivity and supplementary analyses performed under different data handling strategies for missing data and intercurrent events were consistent with the primary efficacy results, leading to the same conclusion for a robust interpretation of the noninferiority findings ([Table 22](#)).

Additionally, PHOTON study results for the key secondary endpoint of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS at Week 48 established noninferiority of HDq12 to 2q8 but not HDq16 to 2q8 using a noninferiority margin of -10% ([Table 24](#)). For example, at Week 48, 27%, 29%, and 20% of subjects in the 2q8, HDq12, and HDq16 treatment groups, respectively, achieved ≥ 2 -step improvement in DRSS from baseline to Week 48 with adjusted treatment differences of 2.0% (95% CI: -6.4%, 10.4%) for HDq12 versus 2q8 and -7.5% (95% CI: -16.8%, 1.8%) for HDq16 versus 2q8.

In summary, based on the collective efficacy evidence from the adequate and well controlled trial of PHOTON study, the reviewer concludes that the application for the DME indication provided substantial evidence for comparable efficacy benefit (in BCVA gain) of aflibercept 8 mg administered every 12 weeks and every 16 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after five initial monthly doses.

For the DR indication, the application provided evidence for comparable efficacy benefit of aflibercept 8 mg administered every 12 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after five initial doses. For this indication, the application did not provide evidence of comparable efficacy benefit of aflibercept 8 mg administered every 16 weeks after three initial doses compared to aflibercept 2 mg administered every 8 weeks after five initial doses.

2. INTRODUCTION

2.1. OVERVIEW

In this BLA submission, the Applicant seeks approval of high dose (8 mg) aflibercept for the treatment of subjects with nAMD, DME, and DR.

nAMD is a chronic eye disease generally caused by abnormal growth and leakage of blood vessels in the macula (the central part of the retina responsible for detailed vision); it is a major health issue in aging populations globally, causing greater impact in elderly subjects. DR is the most common microvascular complication of diabetes mellitus; it occurs when the glucose damages the blood vessels in the retina. DME is a complication of DR, and it can develop at any stage of DR when the macula swells; it is the primary cause of vision loss and blindness in subjects with diabetes.

Currently, anti-vascular endothelial growth factor (VEGF) injections are the standard of care for treatment of these eye diseases. Although many subjects benefit from the treatment, some need the injections as frequently as every 4 to 8 weeks, especially in the first year of treatment. Aflibercept (EYLEA®) was approved for the treatment of these diseases with 2 mg (administered at a concentration of 40 mg/mL with an intravitreal injection) every 4 to 8 weeks after three initial monthly injections (nAMD) and five initial monthly injections (DME/DR).

In this BLA, the Applicant seeks approval of aflibercept administered in a higher dose (8 mg, administered at a concentration of 114.3 mg/mL) in a dosing interval of 12 weeks or longer for the mentioned indications.

Prior to the current submission, the clinical development program for HD aflibercept was discussed with the Agency on several occasions under pre-investigational new drug (PIND) 141278 and investigational new drug 12462:

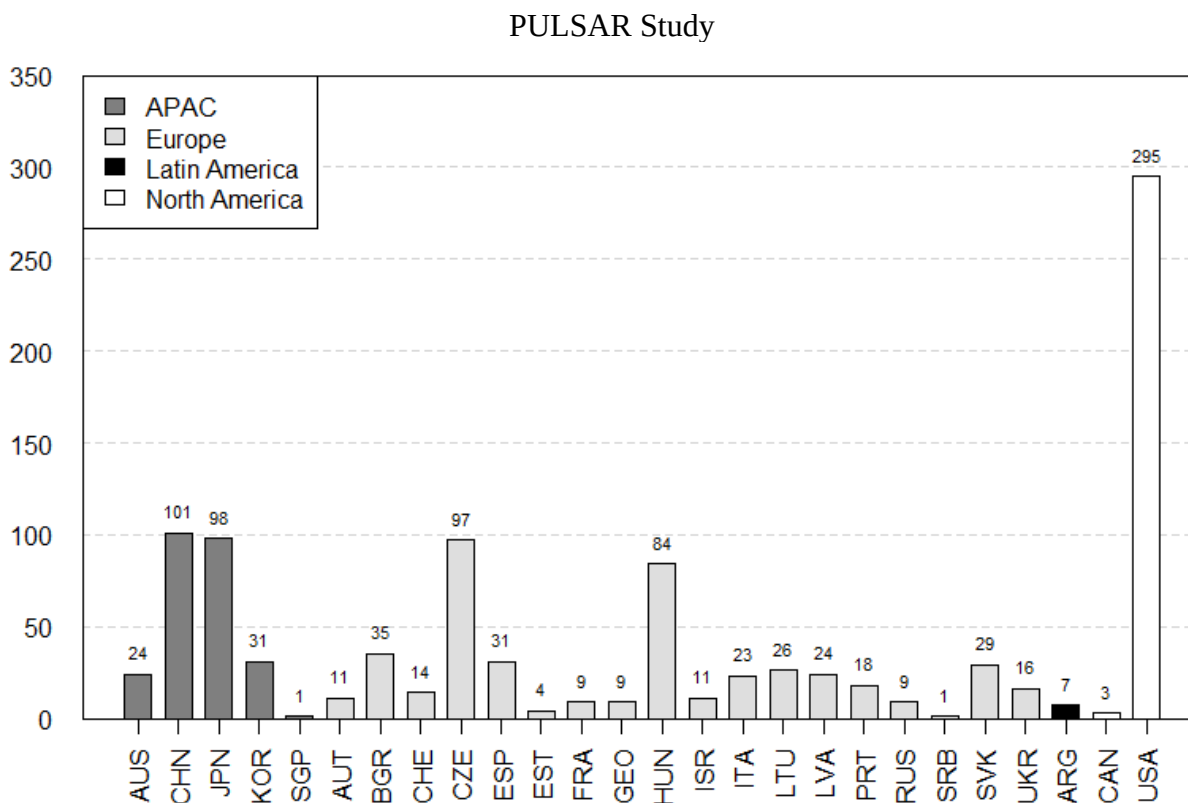
- i) A Type B Pre-IND meeting was held on January 22, 2019, and the adequacy of the proposed clinical development plan for the indications was discussed, including whether superiority or noninferiority studies are appropriate.
- ii) A follow up email from Agency, sent on February 26, 2019, confirmed that a single noninferiority study for nAMD and a single noninferiority study for DME are acceptable for the submission, with extended dosing intervals.
- iii) A Type C teleconference meeting was held on November 29, 2021, where the Agency agreed that it is acceptable to submit an original BLA to register HD aflibercept for the indications of nAMD and DME/DR.
- iv) A Type B Pre-BLA teleconference meeting was held on September 26, 2022, and the study protocols for nAMD (PULSAR) and DME/DR (PHOTON) were assessed. Among other comments, the Agency stressed that a noninferiority margin of 15% for the secondary endpoint of the proportion of subjects who achieve ≥ 2 -step improvement in the Diabetic Retinopathy Severity Scale (DRSS) from baseline at Week 48, for the indication of DR in the PHOTON study is not acceptable. A noninferiority margin of 10% was recommended.

2.1.1. Specific Studies Reviewed

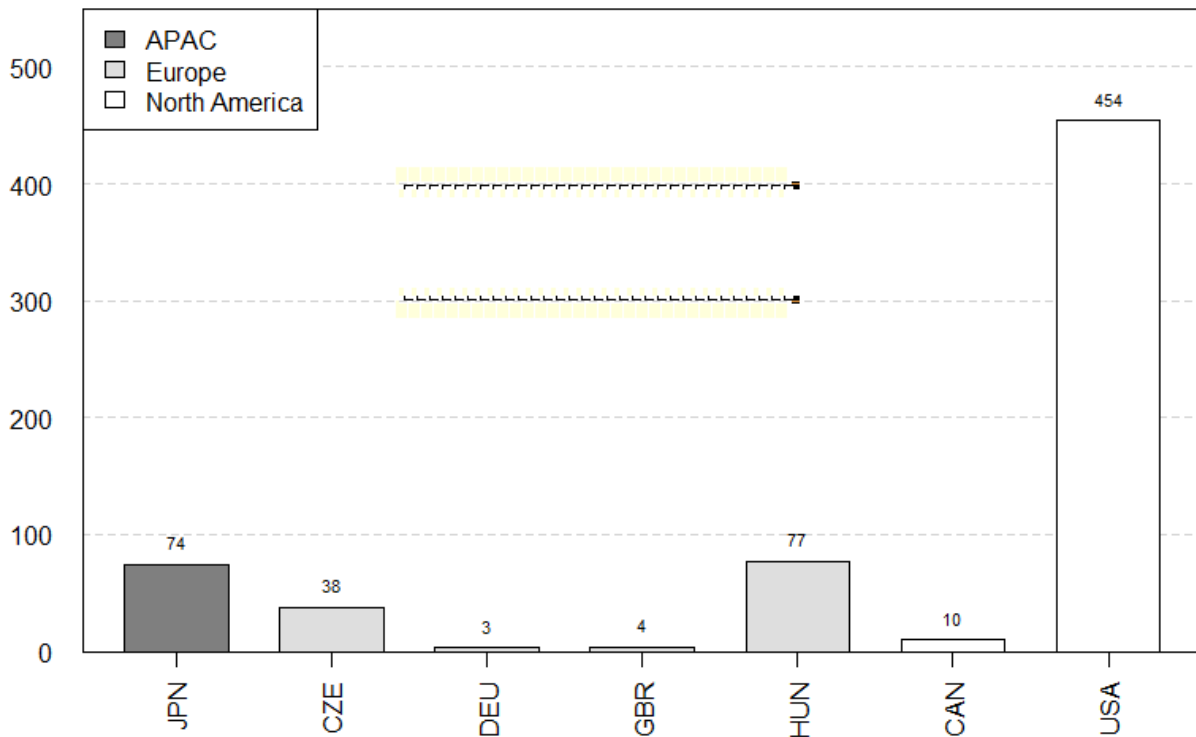
The BLA submission included data from two ongoing pivotal Phase 3 studies – one study designed to support the nAMD indication (PULSAR [20968]) and one study designed to support the DME/DR indications (PHOTON [VGFTe-HD-DME-1934]).

In PULSAR and PHOTON studies, respectively, a total of 1011 and 660 subjects were enrolled globally. [Figure 1](#) displays the number of subjects enrolled by region in the two studies. As shown, most subjects enrolled in these studies were from USA. For more details on the design of the PULSAR study for the nAMD indication and on the design of the PHOTON study for the DME/DR indications, see [Section 3.2.1](#) and [Section 3.2.2](#), respectively.

Figure 1: Number of Subjects Enrolled by Region. Top panel: PULSAR study. Bottom panel: PHOTON study.



PHOTON Study



Source: Based on Reviewer's analysis.

2.2. DATA SOURCES

The data source for this review included the clinical study reports, the analysis and tabulation datasets, study protocols including amendments, corresponding statistical analysis plans, and the integrated summary of safety and efficacy datasets. These are provided in an electronic submission located at <\\CDSESUB1\evsprod\BLA761355\0001>. The primary analysis datasets are located at <\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets>.

The data analyzed in this review are based on the phase 3 study PULSAR [20968] for the nAMD indication and on the phase 2/3 study PHOTON [VGFTe-HD-DME-1934]) for the DME/DR indications.

3. STATISTICAL EVALUATION

3.1. DATA SOURCES

The reviewer found the quality and integrity of the submitted data and analysis acceptable.

3.2. EVALUATION OF EFFICACY

In this section, the efficacy assessments for the PULSAR study to support the nAMD indication are discussed in [Section 3.2.1](#) and for the PHOTON study to support the DME/DR indications are discussed in [Section 3.2.2](#). Within these sections, descriptions of the study designs and the efficacy endpoints, the statistical methodologies used, the summary of patient disposition and demographic and baseline characteristics, and the efficacy results are presented and discussed.

3.2.1. **PULSAR Study for nAMD Indication**

3.2.1.1. *Study Design and Endpoints*

Study Design

Data intended to support efficacy and safety for HD aflibercept (8 mg) versus aflibercept 2 mg for the treatment of nAMD is based on an ongoing phase 3, multi-center, randomized, double-masked, active-controlled noninferiority (NI) global study: the PULSAR study.

The primary study objective was to determine if treatment with HD aflibercept at intervals of 12 or 16 weeks provided noninferior BCVA compared to aflibercept 2 mg dosed every 8 weeks. Of note, aflibercept administered every 8-week interval after three initial monthly injections was approved in the United States and other countries since 2011 for the treatment of nAMD.

In the PULSAR study, a total of 1011 treatment-naïve subjects, at least 50 years of age, with choroidal neovascularization (CNV) secondary to nAMD who had baseline BCVA letter score of 78 to 24 were enrolled.

Eligible subjects who met all the study enrollment criteria were randomized in a 1:1:1 ratio and were to receive aflibercept 2 mg every 8 weeks (2q8, n=337), aflibercept 8 mg every 12 weeks (HDq12, n=336), or aflibercept 8 mg every 16 weeks (HDq16, n=338). Subjects were enrolled in 223 sites in 27 countries/regions (Europe [N = 451], North America [N = 298], Latin America (Argentina) [N = 7], Australia [N = 24], and Asia Pacific [N = 231]). Randomization was stratified by baseline BCVA (<60 vs. ≥60) and geographical region (Japan vs. Rest of the World).

The total treatment duration of the PULSAR study is approximately 96 weeks. During this period, subjects assigned in the 2q8 arm were to receive aflibercept 2 mg every 8 weeks following three initial monthly doses, and subjects assigned in the HDq12 and HDq16 arms were to receive aflibercept 8 mg every 12 and every 16 weeks, respectively, following three initial monthly doses. For masking purposes, sham injections were given at all visits where an active injection was not planned. During the 96-week treatment period, all subjects were followed every four weeks through Week 96.

Although the total study duration of PULSAR study is 96 weeks, this BLA submission was based on the first 48 weeks of efficacy and safety data - the remaining portion of the study is still

ongoing. PULSAR was initiated on April 24, 2020, and the data cut-off date for this submission was on August 24, 2022.

In the PULSAR study, dosing intervals for subjects randomized in the HDq12 or HDq16 arms were shortened during the first 48 weeks starting at Week 16 based on protocol-defined dose regimen modifications (DRM) criteria (Appendix 2, [Table 48](#)):

- Subjects in the HDq12 or the HDq16 arm that met the DRM criteria at Week 16 or Week 20 received 8 mg aflibercept at that visit and every 8 weeks thereafter.
- Subjects in the HDq16 arm who did not meet the DRM criteria at Week 16 or 20 but met the criteria at Week 24 received 8 mg aflibercept at that visit and every 12 weeks thereafter.
- Subjects whose interval was not shortened to every 8-week dosing at or before Week 24, the interval was shortened if the DRM criteria was met at a subsequent dosing visit.
 - Subjects in the HDq12 arm who met the DRM criteria received the planned dose at that visit and every 8 weeks thereafter.
 - Subjects in the HDq16 arm who met the DRM criteria received the planned dose at that visit and continued to be dosed every 12 weeks if they were on a 16-week interval or switched to every 8 weeks if they were previously shortened to a 12-week interval.

[Figure 2](#) displays the dosing schedule for PULSAR study.

Figure 2: Dosing schedule for PULSAR study

	Day 1	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24	Wk 28	Wk 32	Wk 36	Wk 40	Wk 44	Wk 48
2q8	X	X	X		X	o	X	o	X	o	X	o	X
HDq12	X	X	X		<u>o^a</u>	<u>X^a</u>	o	o	<u>X^c</u>	o	o	<u>X^c</u>	o
HDq16	X	X	X		<u>o^b</u>	<u>o^b</u>	<u>X^b</u>	o	o	o	<u>X^c</u>	o	o

	Wk 52	Wk 56	Wk 60	Wk 64	Wk 68	Wk 72	Wk 76	Wk 80	Wk 84	Wk 88	Wk 92	Wk 96
2q8	o	X	o	X	o	X	o	X	o	X	o	
HDq12	o	<u>X^d</u>	o	o	<u>X^d</u>	o	o	<u>X^d</u>	o	o	<u>X^d</u>	
HDq16	o	<u>X^d</u>	o	o	o	<u>X^d</u>	o	o	o	<u>X^d</u>	o	

1° = primary, 2° = secondary, DRM = dose regimen modification, HD = high dose, IXRS = Interactive Response System, q8 = every 8 weeks, q12 = every 12 weeks, q16 = every 16 weeks, Wk = Week

X = active injection, o = sham procedure

For masking purposes, DRM assessments were performed in all participants at all visits (through the IXRS) starting from Week 16.

a HDq12 group: If DRM criteria were met at Week 16 or 20, participants continued on q8 rescue regimen.

b HDq16 group: If DRM criteria were met at Week 16 or 20, participant continued on q8 rescue regimen. If DRM criteria were met at Week 24, participant continued on q12 regimen.

c For participants remaining on a dosing interval of q12 or q16 weeks after Week 24, if DRM criteria are met at an active injection visit, the next dosing interval will be reduced by 4 weeks (to a minimum of q8).

d From Week 52, all participants in HD groups will be eligible for dose interval shortening (to a minimum of q8) or extension (by 4-week increments) according to pre-specified DRM criteria. If DRM criteria are met at an active injection visit, the next dosing interval will be changed by 4 weeks.

Source: Figure 7-2 of PULSAR Clinical Study Report.

Efficacy Evaluation

Key efficacy evaluation in PULSAR study was based on (i) BCVA letter score as measured using the Early Treatment of Diabetic Retinopathy Study (ETDRS) protocol at 4 meters at each study visit, and (ii) the presence or absence of intraretinal fluid (IRF) and subretinal fluid (SRF) measured using spectral domain optical coherence tomography (SD-OCT).

BCVA as well as IRF and SRF were measured in the study eye every 4 weeks through Week 48.

Study Endpoints

The primary and key secondary efficacy endpoints were *the change from baseline in BCVA at Week 48 and the proportion of subjects with no IRF and no SRF in central subfield at Week 16*, respectively. Additional secondary efficacy endpoints were:

- Proportion of subjects gaining ≥ 15 letters in BCVA at Week 48.
- Proportion of participants with no IRF and no SRF in the central subfield at Week 48.
- Change from baseline in central subfield retinal thickness (CST) at Week 48.
- Proportion of subjects with BCVA ≥ 69 letters at Week 48.
- Change in choroidal neovascularization (CNV) size from baseline to Week 48.
- Change in total lesion area from baseline to Week 48.
- Change from baseline in National Eye Institute Visual Functioning Questionnaire-25 (NEI-VFQ-25) total score at Week 48.

3.2.1.2. Statistical Methodologies

The primary analysis was based on the estimand concept. The estimand of primary interest was based on a combination of hypothetical and treatment policy strategies. It described the change from baseline for all subjects who started treatment, under a pre-specified data handling strategy for occurrence of intercurrent events (see detail below).

Sample Size Determination

The Applicant planned to enroll 960 subjects (320 subjects per arm). The sample size calculation was based on demonstrating noninferiority of HDq16 and HDq12 to 2q8 in the primary efficacy variable of the change in BCVA from baseline at Week 48. A sample size of 320 subjects per arm provided, under the hierarchical testing procedure, at least 94% power for comparing HDq12 vs. 2q8. The power to reject both primary null hypotheses (HDq12 vs. 2q8 and HDq16 vs. 2q8) was 88%, using a noninferiority margin of -4 letters.

The sample size calculation was based on a 1-sided t-test assuming a true treatment difference of zero and a common standard deviation of 13.6 letters for the mean change in BCVA from baseline at Week 48, with a one-sided significance level of 2.5% and a dropout rate of 10%.

Analysis Populations

Three analysis populations were defined for the analyses of the efficacy and safety variables in PULSAR study: (i) full analysis set (FAS) - included all randomized subjects who received at least one dose of study medication; it is based on the treatment assigned to the subjects at baseline (as randomized), (ii) per-protocol set (PPS) - included all subjects in the FAS who had a baseline and at least one post-baseline assessment of BCVA, and did not have any relevant important protocol deviation that affected the primary efficacy variable, and (iii) safety analysis set (SAF) - included all randomized subjects who received any study treatment; it is based on the treatment received (as treated).

The primary analysis set for efficacy assessment was the FAS. Supplementary analysis based on the PPS was considered as supporting.

Justification of the Noninferiority Margin

The primary efficacy objective of the study was to demonstrate noninferiority of HDq16 and HDq12 to 2q8 in the mean change in BCVA from baseline at Week 48 using a noninferiority margin of -4.0 letters. In the study protocol and corresponding statistical analysis plan, the Applicant provided a reasonable justification for the choice of the NI margin.

Reviewer's Comment:

Based on our understanding of the efficacy data from previous studies with anti-VEGF therapies (e.g., Lucentis [ranibizumab] and Eylea [aflibercept]), the Applicant's proposed noninferiority margin of -4 letters for the mean change in BCVA from baseline at Week 48 was acceptable.

Primary Estimand

The Applicant's primary estimand of interest was the difference in the least square mean change in BCVA from baseline to Week 48 between each high dose of aflibercept (HDq12 and HDq16) versus aflibercept 2 mg (2q8), for all randomized subjects in FAS under the following data handling strategies for those who had occurrence of intercurrent events (ICEs):

- Data for subjects who (i) discontinued study intervention but stayed in the study, (ii) missed an active injection and did not get it at next visit, (iii) used a prohibited medication after baseline visit, or (iv) received a sham injection instead of an active were censored and assumed that these subjects adhered to the treatment under a hypothetical strategy.
- Data for subjects who discontinued study intervention and study at the same time was also assumed that these subjects adhered to the treatment under a hypothetical strategy.
- Data for subjects who (i) missed an active injection and got it at next visit, (ii) received an active injection instead of sham, or (iii) received a wrong dose level were used regardless of the occurrence of ICE under a treatment policy strategy.

[Table 3](#) displays the summary of subjects with occurrence of ICEs through Week 48 in the PULSAR study.

Table 3: Summary of Intercurrent Events Through Week 48 in PULSAR study

ICE Category	2q8	HDq12	HDq16
Premature discontinuation of study intervention but continued study	6 (1.8%)	4 (1.2%)	4 (1.2%)
Missed study intervention was planned to be an active injection, but did not take make-up injection at next visit	16 (4.8%)	11 (3.3%)	15 (4.4%)
Use of a prohibited medication after the first dose of study medication by Week 48, inclusive	2 (0.6%)	0	1 (0.3%)

Source: Post-text Table 14.2.1 of PULSAR Study.

Reviewer's Comment:

In the PULSAR study, few subjects experienced intercurrent events during the 48-week treatment period ([Table 3](#)). Thus, the impact of the data handling strategies on the overall conclusion expected to be very minimal.

Primary Efficacy Analysis for Primary Estimand

The primary efficacy analysis was an evaluation of noninferiority of each high dose of aflibercept (HDq12 and HDq16) to 2q8 in the primary efficacy variable using a mixed model for repeated measures (MMRM) analysis based on the FAS. The model included baseline BCVA measurement as a covariate and treatment group, baseline BCVA category, geographical region, and visit as fixed effects as well as visit-by-treatment group and visit-by-baseline BCVA interactions. A Kenward-Roger approximation was used for the denominator degrees of freedom. The model assumed an unstructured covariance structure to account for within-subject correlation. Only data up to Week 48 was used in this analysis.

Based on the MMRM method, HDq12 arm or HDq16 arm was considered noninferior to the active control 2q8 arm if the lower limit of the two-sided 95% confidence interval (CI) estimate for the treatment difference in the least square means of the change in BCVA from baseline to Week 48 (high dose minus 2q8) was greater than -4.0 letters.

Superiority of each high dose of aflibercept to 2q8 in the primary efficacy endpoint was also assessed if (i) noninferiority of each high dose of aflibercept to 2q8 in the primary efficacy endpoint and (ii) superiority of pooled high dose vs. 2q8 in the key secondary endpoint were established.

Sensitivity Analyses to the Primary Estimand

The Applicant performed the following sensitivity analyses for the primary estimand to assess the robustness of the primary efficacy analysis in the presence of missing data or occurrence of ICEs:

- Last observation carried forward (LOCF) approach was conducted for subjects who had at least one post-baseline value but discontinued prematurely or had an ICE before Week 48 and ANCOVA was applied for the change from baseline in BCVA at Week 48.
- Multiple imputation (MI) approach assuming a missing at random (MAR) missing data mechanism was implemented and ANCOVA was applied for the change from baseline in BCVA at Week 48. A tipping point analysis was performed to assess the validity of the MAR assumption.

The primary efficacy analysis was repeated on the PPS as a supporting analysis.

Secondary Efficacy Analysis

The secondary efficacy analysis was an evaluation of superiority of the pooled high dose (HDq12 and HDq16) to 2q8 in the key secondary endpoint of the proportion of subjects with no IRF and no SRF in central subfield at Week 16 based on a stratified Cochran-Mantel-Haenszel (CMH) test, stratified by baseline BCVA category and geographical region.

Based on the CMH test, a weighted point and two-sided 95% CI estimates for the difference in proportions (pooled high dose minus 2q8) using the CMH weights and normal approximation of the weighted estimates was provided. Based on the two-sided 95% CI estimates, the pooled high dose of aflibercept was considered superior to 2q8 if the lower limit of the CI was greater than 0%.

The Applicant's secondary efficacy analysis was based on the FAS using LOCF approach for handling missing data.

All other binary secondary efficacy variables were analyzed similarly to the key secondary variable and all continuous secondary variables were analyzed similarly to the primary efficacy variable.

Reviewer's Comment:

For the evaluation of the key secondary endpoint at Week 16, pooling of the two high doses of aflibercept was made because subjects assigned in both arms received the same number of injections prior up to Week 16.

Type I Error Control (Plan for Multiplicity Adjustment)

For testing noninferiority followed by superiority of each high dose of aflibercept to 2q8 in the primary and key secondary efficacy endpoints, the Applicant implemented the following hierarchical testing strategy to control the study-wise Type I error rate at two-sided significance level of 5%:

1. Noninferiority of HDq12 to 2q8 in the mean change in BCVA at Week 48.

2. Noninferiority of HDq16 to 2q8 in the mean change in BCVA at Week 48.
3. Superiority of pooled high dose (HDq12 and HDq16) vs. 2q8 in the proportion of subjects with no IRF and no SRF in central subfield at Week 16.
4. Superiority of HDq12 to 2q8 in the mean change in BCVA at Week 48.
5. Superiority of HDq16 to 2q8 in the mean change in BCVA at Week 48.

Subgroup Analysis

Subgroup analyses of the primary efficacy endpoint were performed by the following subgroup variables: sex, age (<75 years vs. ≥75 years), race (White vs. Asian vs. Others), baseline BCVA category (<60 vs. ≥60 letters), and region (Japan vs. Rest of the World).

3.2.1.3. Patient Disposition, Demographic and Baseline Characteristics

Patient Disposition

Table 4 shows the summary of subject disposition and the primary reasons for study discontinuation during the 48-week treatment period in the PULSAR study. Overall, 1011 subjects were randomized. Most subjects completed the 48-week treatment period (about 93%).

A total of 68 (7%) subjects discontinued from the study (7% in 2q8 arm, 5% in HDq12, and 7% in HDq16). Discontinuation rates were comparable across the treatment groups. The most common reason for discontinuation was withdrawal by subject (2.2%) followed by adverse event (1.1%) and death (0.9%). In PULSAR study, a total of 9 subjects (5 in 2q8, 3 in HDq12, and 1 in HDq16) died during the 48-week treatment period.

Table 4: Summary of Subject Disposition and Reasons for Study Discontinuation
(All Randomized Subjects)

	2q8	HDq12	HDq16	Total
Randomized, n (%)	337 (100%)	336 (100%)	338 (100%)	1011 (100%)
Treated, n (%)	336 (99.7%)	335 (99.7%)	338 (100%)	1009
Completed study at Week 48, n (%)	309 (91.7%)	316 (94.0%)	312 (92.3%)	937 (92.7%)
Unknown, n (%)	3 (0.9%)	2 (0.6%)	1 (0.3%)	6 (0.6%)
Discontinued from Study	25 (7.4%)	18 (5.4%)	25 (7.4%)	68 (6.7%)
Primary Reason for Discontinuation				
Adverse Event	5 (1.5%)	1 (0.3%)	5 (1.5%)	11 (1.1%)
Death	5 (1.5%)	3 (0.9%)	1 (0.3%)	9 (0.9%)
Lack of Efficacy	2 (0.6%)	0	0	2 (0.2%)
Lost to Follow-Up	1 (0.3%)	1 (0.3%)	0	2 (0.2%)
Physician Decision	1 (0.3%)	3 (0.9%)	2 (0.6%)	6 (0.6%)
Protocol Deviation	0	1 (0.3%)	1 (0.3%)	2 (0.2%)
Subject Decision: Covid-19 related	2 (0.6%)	2 (0.6%)	2 (0.6%)	6 (0.6%)
Withdrawal by Subject	5 (1.5%)	5 (1.5%)	12 (3.6%)	22 (2.2%)
Other	4 (1.2%)	2 (0.6%)	2 (0.6%)	8 (0.8%)

Source: Table 8.1 of PULSAR Clinical Study Report and Reviewer's analysis.

Analysis Population

Table 5 displays the number of subjects by analysis populations. A total of 1011 subjects were randomized. Of these, two subjects were removed from the FAS and SAF because they did not take the study intervention. A total of 41 subjects were excluded from the PPS, and a total of 270 (98 in 2q8, 87 in HDq12, and 85 in HDq16) had major protocol deviations.

Table 5: Summary of Analysis Populations in PULSAR Study

	2q8	HDq12	HDq16	Total
Randomized subjects	337	336	338	1011
Full Analysis Set	336	335	338	1009
Safety Analysis Set (As Treated)	336	335	338	1009
Per-Protocol Set	320	325	325	970
Major protocol Deviations	98	87	85	270
Major protocol Deviations due to COVID-19	14	11	10	35

Source: Table 8-3 of PULSAR Clinical Study Report and Reviewer's analysis.

Demographic and Baseline Characteristics

The summary of the demographic and baseline characteristics for subjects in the FAS in the PULSAR study are presented in Table 6. Most subjects were female (55%) and white (76%). The average age was about 75 years (range: 50 to 96 years) with higher percentage in the age group 65-75 years (38%). Most subjects were from Europe (45%) and North America (30%).

The average baseline BCVA in the study eye was 59.6 letters with a majority (58%) of subjects had a baseline BCVA of ≥ 60 letters.

In general, all demographic and baseline characteristics were well balanced across the three treatment groups.

Table 6: Demographic and Baseline Characteristics in PULSAR Study (Full Analysis Set)

	2q8 (N = 336)	HDq12 (N = 335)	HDq16 (N = 338)	Total (N = 1009)
Age (years)				
Mean (SD)	74.2 (8.8)	74.7 (7.9)	74.5 (8.5)	74.5 (8.4)
Median	74	75	75	75
Range	50 - 96	52 - 93	53 - 95	50 - 96
Age Category				
< 65	45 (13.4%)	26 (7.8%)	43 (12.7%)	114 (11.3%)
65-<75	126 (37.5%)	137 (40.9%)	124 (36.7%)	387 (38.4%)
75-<80	69 (20.5%)	77 (23.0%)	66 (19.5%)	212 (21.0%)
80-<85	59 (17.6%)	59 (17.6%)	66 (19.5%)	184 (18.2%)
≥ 85	37 (11.0%)	36 (10.7%)	39 (11.5%)	112 (11.1%)
Sex				
Female	188 (56.0%)	182 (54.3%)	180 (53.3%)	550 (54.5%)
Male	148 (44.0%)	153 (45.7%)	158 (46.7%)	459 (45.5%)

	2q8 (N = 336)	HDq12 (N = 335)	HDq16 (N = 338)	Total (N = 1009)
Race				
Asian	83 (24.7%)	74 (22.1%)	77 (22.8%)	234 (23.2%)
Black or African American	2 (0.6%)	2 (0.6%)	0	4 (0.4%)
Multiple	0	1 (0.3%)	0	1 (0.1%)
White	249 (74.1%)	256 (76.4%)	260 (76.9%)	765 (75.8%)
Not Reported	2 (0.6%)	2 (0.6%)	1 (0.3%)	5 (0.5%)
Ethnicity				
Hispanic or Latino	12 (3.6%)	7 (2.1%)	9 (2.7%)	28 (2.8%)
Not Hispanic or Latino	322 (95.8%)	322 (96.1%)	326 (96.4%)	970 (96.1%)
Not Reported	2 (0.6%)	6 (1.8%)	3 (0.9%)	11 (1.1%)
Region 1				
Asia Pacific	89 (26.5%)	80 (23.9%)	84 (24.9%)	253 (25.1%)
Europe	137 (40.8%)	162 (48.4%)	152 (45.0%)	451 (44.7%)
Latin America (Argentina)	2 (0.6%)	1 (0.3%)	4 (1.2%)	7 (0.7%)
North America (US and Canada)	108 (32.1%)	92 (27.5%)	98 (29.0%)	298 (29.5%)
Region 2				
Japan	33 (9.8%)	31 (9.3%)	33 (9.8%)	97 (9.6%)
Rest of the World	303 (90.2%)	304 (90.7%)	305 (90.2%)	912 (90.4%)
Baseline BCVA (letters)				
Mean (SD)	58.9 (14.0)	59.9 (13.4)	60.0 (12.4)	59.6 (13.3)
Median	62.0	62.0	61.0	62.0
Range	24 - 78	24 - 78	24 - 78	24 - 78
Baseline BCVA Category				
≥ 60	200 (59.5%)	194 (57.9%)	194 (57.4%)	588 (58.3%)
< 60	136 (40.5%)	141 (42.1%)	144 (42.6%)	421 (41.7%)
Baseline Intraocular Pressure				
Mean (SD)	14.8 (3.0)	14.9 (3.2)	14.9 (3.2)	14.9 (3.2)
Median	15.0	15.0	15.0	15.0
Range	6 - 25	7 - 25	7 - 25	6 - 25
Baseline Central Subfield Retinal Thickness				
Mean (SD)	367.1 (133.6)	370.3 (123.7)	370.7 (132.7)	369.3 (130.0)
Median	343.0	348.0	340.0	344.0
Range	142 – 1116	151 – 840	144 – 913	142 – 1116

Source: Table 8-5 and Table 8-6 of PULSAR Clinical Study Report.

3.2.1.4. Results and Conclusions

Primary Efficacy Endpoint: Change in BCVA from Baseline at Week 48

The primary efficacy objective of the PULSAR study was an assessment of noninferiority of each high dose of aflibercept (HDq12 and HDq16) to aflibercept 2q8 in the primary efficacy endpoint in the FAS. The primary estimand of interest to address this objective was the difference in least square mean change in BCVA from baseline at Week 48.

Table 7 below shows the number of subjects with observed BCVA data in the PULSAR study at each visit. As stated in Section 3.2.1.2 (*RE: Primary Estimand*), data under the hypothetical strategy were censored and data under the treatment policy strategy were included in the summary table.

Table 7: Number of Subjects with Observed BCVA Data by Visit (PULSAR)

Treatment	Visit (in Weeks)												
	0	4	8	12	16	20	24	28	32	36	40	44	48
2q8	336	335	329	327	326	314	315	305	303	293	293	279	285
HDq12	335	332	331	328	325	322	319	317	314	304	300	298	299
HDq16	337	337	332	330	328	324	316	311	306	302	296	291	289

Source: Based on Reviewer's analysis.

Based on the above observed data in the combined hypothetical and treatment policy strategy, Table 8 shows the summary of the mean change in BCVA from baseline to Week 48 and the adjusted mean estimates for each treatment group at this visit and the treatment differences in the adjusted means from the MMRM method.

Table 8: Summary of Mean Change in BCVA from Baseline at Week 48 (FAS) (PULSAR)

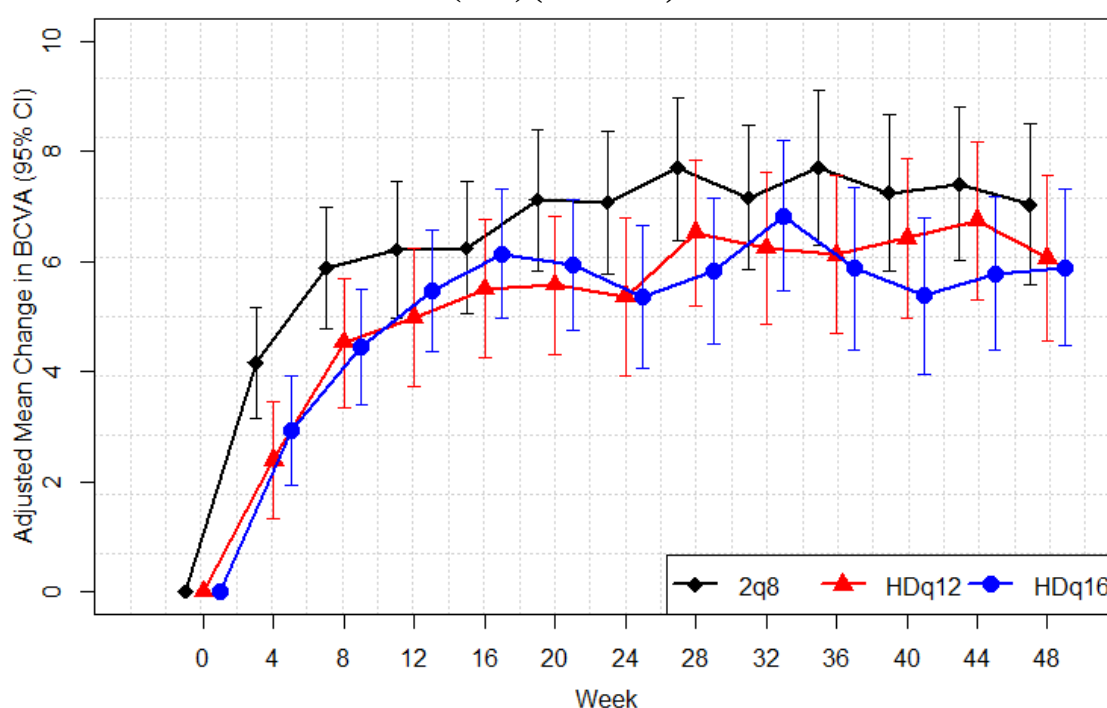
Analysis Visit		2q8	HDq12	HDq16
Mean BCVA				
Baseline	Mean (SD)	59.0 (14.01)	59.8 (13.42)	60.1 (12.40)
	Median (Range)	62 (26 – 78)	62 (24 – 78)	61 (24 – 78)
Week 48	Mean (SD)	66.5 (16.58)	66.9 (15.47)	66.3 (15.37)
	Median (Range)	71 (10 - 94)	71 (2 - 97)	70 (9 - 95)
Mean Change in BCVA				
Week 48	Mean (SD)	7.6 (12.20)	6.7 (12.60)	6.2 (11.69)
	Median (Range)	8 (-41 – 49)	6 (-70 – 45)	7 (-43 – 37)
Adjusted Mean Change in BCVA				
Week 48	LS Mean (95% CI)	7.0 (5.6, 8.5)	6.1 (4.5, 7.6)	5.9 (4.5, 7.3)
	Diff (95% CI)	---	-1.0 (-2.9, 0.9)	-1.1 (-3.0, 0.7)

Source: Reviewer's analysis based on ADEFF.xpt dataset located at [\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

As shown, subjects treated with high doses of aflibercept (HDq12 or HDq16) had a noninferior mean change in BCVA from baseline at Week 48 compared to subjects treated with aflibercept 2 mg (2q8) because the 95% lower confidence limits for the differences in the adjusted means between each high dose of aflibercept to aflibercept 2 mg was greater than the noninferiority margin of -4.0 letters. The adjusted mean change in BCVA from baseline to Week 48 was +7.0 letters in 2q8 group, +6.1 letters in HDq12 group, and +5.9 letters in HDq16 group with a treatment difference of -1.0 (95% CI: -2.9 to 0.9) between HDq12 and 2q8 and -1.1 (95% CI: -3.0 to 0.7) between HDq16 and 2q8.

Figure 3 display a plot of the adjusted mean changes in BCVA from baseline to each visit through Week 48 including the 95% confidence interval estimates (vertical bars at each visit).

Figure 3: Plot of Adjusted Mean Change in BCVA from Baseline Through Week 48 (FAS) (PULSAR)



Source: Based on Reviewer's analysis (Appendix 1, Table 35)

As shown, the adjusted mean changes from baseline at each visit through Week 48 appear to be numerically higher in the 2q8 group compared to in both high dose groups. However, the confidence intervals at each visit overlapped between the treatment groups implying a lack of treatment difference. The numerically higher adjusted mean changes that subjects in the 2q8 group exhibited compared to subjects in the high dose groups could be attributed to the frequent injections that subjects in the 2q8 arm received through Week 48.

Reviewer's Remark:

The wave-like pattern in Figure 3 in each treatment group after Week 12 is explained by the extended dosing regimen. For example, for subjects in the 2q8 arm, the BCVA gain was at the

lowest point (troughs) at visits where the subject received injection (i.e., at Weeks 16, 24, 32, 40, and 48) and at the highest point (peaks) at visits immediately followed by the injection visit (i.e., at Weeks 20, 28, 36, and 44). The same but extended pattern was seen in the HDq12 and HDq16 groups. Because the wave-like pattern was dependent on the timing of the dose, BCVA assessment at Week 48 alone may benefit one arm over the other depending on when the last injection was received at the prior visits. Thus, to evaluate the impact of this concern, the reviewer compared the treatment groups based on the change in BCVA data from baseline averaged at Week 40, 44, and 48 (Table 9). During these three visits, subjects in each treatment group had one injection (at Week 40 for 2q8 and HDq16, and at Week 44 for HDq12). As shown, using this endpoint, the noninferiority conclusion was unaltered which confirm the robustness of the primary efficacy result at Week 48.

Table 9: Summary of Mean Change in BCVA from Baseline at Week 40, 44, 48, and averaged at Week 40/44/48 (FAS) (PULSAR)

Analysis Visit	2q8	HDq12	HDq16
Week 40	7.2 (0.7)	6.4 (0.7)	5.4 (0.7)
Week 44	7.4 (0.7)	6.7 (0.7)	5.8 (0.7)
Week 48	7.0 (0.7)	6.1 (0.8)	5.9 (0.7)
Week 40/44/48	7.5 (0.6)	6.7 (0.7)	6.0 (0.6)
Diff (95% CI)	---	-0.8 (-2.6, 0.9)	-1.5 (-3.3, 0.2)

Source: Based on Reviewer's analysis.

Sensitivity/Supporting Analyses to the Primary Efficacy Endpoint

Different sensitivity and supporting analyses were performed to assess the robustness of the results of the primary efficacy analysis based on the MMRM method. ANCOVA method was used to analyze the primary efficacy endpoint by imputing missing data using LOCF approach and using multiple imputation (MI) approach under MAR assumption (Table 10). Additionally, to check the robustness of the MMRM results if the missing data are not MAR, a tipping point analysis was also performed based on the MI analysis and the results are provided in Appendix 1, Table 36.

Table 10: Sensitivity and Supportive Analyses: Mean Change in BCVA from Baseline at Week 48 (PULSAR)

		2q8	HDq12	HDq16
Sensitivity Analysis: Using ANCOVA method				
LOCF in FAS	LS Mean (95% CI)	7.0 (5.4, 8.7)	6.0 (4.3, 7.6)	5.8 (4.2, 7.5)
	Diff (95% CI)	---	-1.1 (-2.9, 0.8)	-1.2 (-3.1, 0.6)
MI in FAS	LS Mean (95% CI)	7.2 (5.4, 9.0)	6.2 (4.5, 7.9)	6.1 (4.4, 7.8)
	Diff (95% CI)	---	-1.0 (-2.9, 0.9)	-1.1 (-3.0, 0.8)
Supporting Analysis: Using MMRM method				
PPS in FAS	LS Mean (95% CI)	7.1 (5.6, 8.6)	6.1 (4.5, 7.6)	5.9 (4.4, 7.4)
	Diff (95% CI)	---	-1.1 (-3.0, 0.9)	-1.2 (-3.1, 0.7)

LOCF: Last observation carried forward; MI: Multiple imputation; PPS: Per Protocol Set; MMRM: Mixed model for repeated measures; FAS: Full analysis set.

Source: Reviewer's analysis based on ADEFF.xpt dataset located at

[\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

As shown in [Table 10](#) above, the results based on the LOCF, and the MI missing data handling approach were consistent with those of the primary analysis results using MMRM in the FAS. Additionally, the tipping point analysis in [Table 36](#) shows that the smallest shift parameter (delta) used as penalty for subjects with missing data in the HD groups, for which noninferiority could not be shown anymore, was delta = -10 letters for HDq12 vs. 2q8 and delta = -7 letters for HDq16 vs. 2q8. The "tipping point" values of -10 letters and -7 letters requiring such a large difference in BCVA at Week 48 between HDq16 and HDq12 vs. 2q8, respectively, do not seem plausible implying the primary analysis of noninferiority conclusion based on MMRM method is robust to the departure of the MAR assumption.

As supportive analysis, the analysis of the primary efficacy endpoint was repeated using the Per Protocol Set (PPS). These results were consistent with the primary efficacy analysis results ([Table 10](#)). An additional supportive analysis for the primary efficacy endpoint was also conducted based only on subjects who had information at Week 48 and did not have their assigned dosing interval shortened based on the DRM criteria. These results were consistent with the primary efficacy analysis results.

Key Secondary Efficacy Endpoint: Proportion of Participants with no IRF and no SRF at Week 16

This key secondary efficacy endpoint described the proportion of all subjects with no intraretinal fluid (IRF) and no subretinal fluid (SRF) in central subfield at Week 16. Results of the analysis are presented in [Table 11](#), based on LOCF data. For this endpoint, HDq12 and HDq16 arms were pooled since at Week 16 both groups were treated identically with three initial monthly doses.

As noted in [Section 3.2.1.2 \(RE: Secondary Efficacy Analysis\)](#), the binary key secondary endpoint was analyzed using a stratified Cochran-Mantel-Haenszel test ([Table 11](#)). As shown, at Week 16, 63.3% of subjects in the pooled HD groups had no retinal fluid compared to 51.6% of subjects in the 2q8 group with a treatment difference of 11.7% (95% CI: 5.2%, 18.2%). Superiority was shown in this comparison.

Table 11: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 16, LOCF (FAS)

	2q8 (N = 335)	HDq12 (N = 333)	HDq16 (N = 334)	All HD (N = 667)
Participants who had no IRF and no SRF, n (%)	173 (51.6%)	205 (61.6%)	217 (65.0%)	422 (63.3%)
% Difference (95% CI)	--	--	--	11.7 (5.2, 18.2)

LOCF: Last observation carried forward; IRF: intraretinal fluid; SRF: subretinal fluid; Num/Den: Numerator/Denominator.

Source: Reviewer's analysis based on ADEFF.xpt dataset located at

[\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

The key secondary efficacy analysis was repeated on the PPS as supportive analysis and the results were consistent with those obtained in the FAS above (Appendix 1, [Table 37](#)). As a sensitivity analysis, the key secondary analysis was repeated based only on the observed cases in the FAS (Appendix 1, [Table 37](#)). As shown, the results based on the observed data and missing data imputed using the LOCF approach provided similar conclusion except for minor numerical differences.

Secondary Efficacy Endpoint: Proportion of Subjects with no IRF and no SRF at Week 48

The proportion of subjects with no IRF and no SRF at Week 48 was also evaluated as a secondary endpoint. At Week 48, more subjects in HD groups (71.1% in HDq12 and 66.8% in HDq16) had no IRF and no SRF in the center subfield compared to 59.4% of subjects in the 2q8 group with treatment differences of 11.7% (95% CI: 4.6%, 18.9%) for HDq12 vs. 2q8 and 7.5% (95% CI: 0.2%, 14.7%) for HDq16 vs. 2q8 (Table 12).

Table 12: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 48 (FAS) (PULSAR)

	2q8 (N = 335)	HDq12 (N = 332)	HDq16 (N = 334)
Subjects who had no IRF and no SRF, n (%)	199 (59.4%)	236 (71.1%)	223 (66.8%)
Difference % (95% CI)	--	11.7 (4.6, 18.9)	7.5 (0.2, 14.7)

LOCF: Last observation carried forward; IRF: intraretinal fluid; SRF: subretinal fluid; Num/Den: Numerator/Denominator.

Source: Reviewer's analysis based on ADEFF.xpt dataset located at

[\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

A sensitivity analysis performed based on the observed data resulted in a similar conclusion as the one with the LOCF approach except for minor numerical differences (Appendix 1, Table 38).

Reviewer's Comment:

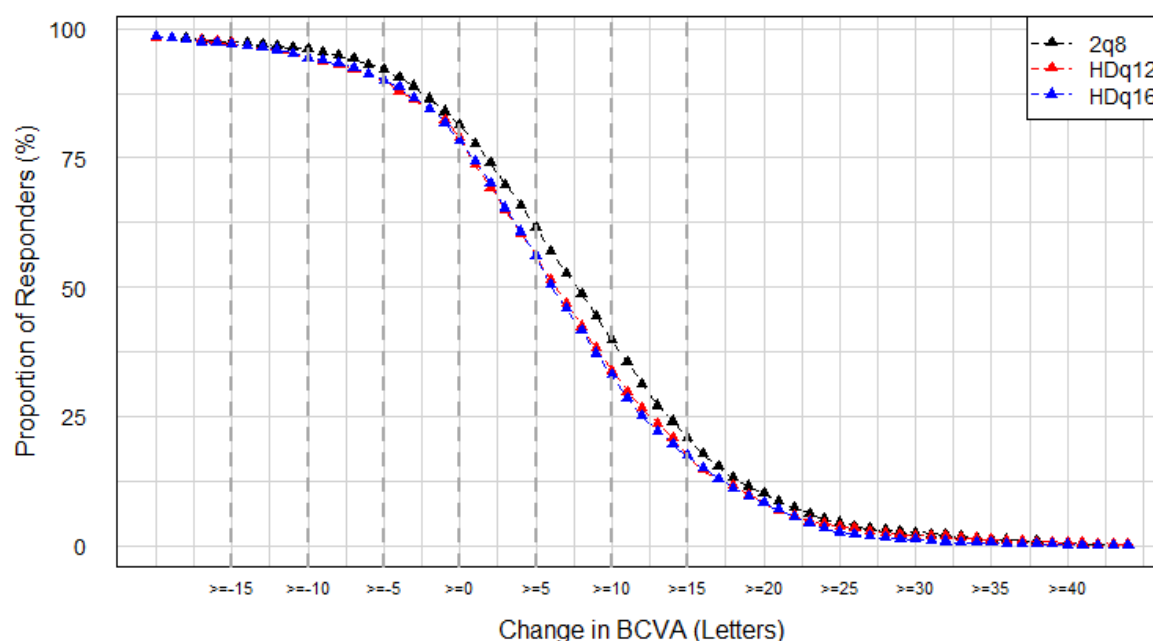
Although the pooled HD groups at Week 16 and the individual HDq12 and HDq16 arms at Week 48 demonstrated superiority in the proportion of subjects with no IRF and no SRF compared to 2q8, consistent efficacy results were not observed at each visit through Week 48 because the treatment groups had variable dosing time after Week 16 and an increasing number of subjects in each treatment group exhibited fluid dryness in the immediate visit after injection was received (Figure 18 and Table 39 in Appendix 1). For example, more subjects in the 2q8 group (68%) than in the HDq12 group (50%) and in the HDq16 group (48%) had no IRF and no SRF at Week 20 with a treatment difference of -18% (95% CI: -26% to -11%) favoring 2q8 compared to HDq12 and -20% (95% CI: -27% to -13%) favoring 2q8 compared to HDq16.

Secondary Efficacy Endpoint: Proportion of Subjects Who Gained/Avoided Losing Letters in BCVA from Baseline at Week 48

The proportion of subjects who gained ≥ 15 letters in BCVA from baseline at Week 48 was a secondary efficacy endpoint in the PULSAR study. Additionally, the proportion of subjects who gained/avoided losing ≥ 5 or ≥ 10 letters from baseline at Week 48 were defined as exploratory endpoints.

Figure 4 shows the cumulative distribution of the change in BCVA at Week 48 based on missing data imputed using multiple imputation. As shown, the three treatment groups seemed comparable across the binary secondary endpoints (dashed lines).

Figure 4: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (FAS) (PULSAR)



Source: Based on Reviewer's analysis. Missing BCVA data were imputed using multiple imputation.

Table 13 shows the summary of the proportion of subjects who gained/avoided losing letters in BCVA at Week 48 based on observed cases and missing data imputed using LOCF approach. As shown, the treatment groups appeared to be comparable in all considered endpoints. For example, using the LOCF approach, 21% of the subjects in the HDq12 group and 22% of the subjects in the HDq16 group gained ≥ 15 letters from baseline at Week 48, compared to 22% in the 2q8 group with treatment differences of -1.8% (95% CI: -7.8% to 4.3%) and -0.9% (-7.0% to 5.1%), respectively. The results based on observed data provided a similar conclusion with minor numerical differences. The same can be determined in the proportion of subjects who avoided losing letters in BCVA. For example, 97% of the subjects in the 2q8 group avoided losing ≥ 15 letters, compared to 95% of subjects in both HDq12 and HDq16 groups with treatment differences of -1.5% (95% CI: -4.4% to 1.5%) and -2.0% (95% CI: -5.1%, 1.1%), respectively.

Table 13: Proportion of Subjects Who Gained ≥ 15 , ≥ 10 , ≥ 5 and ≥ 0 Letters and Avoided Losing ≥ 5 , ≥ 10 and ≥ 15 Letters in BCVA from Baseline at Week 48 (FAS) (PULSAR)

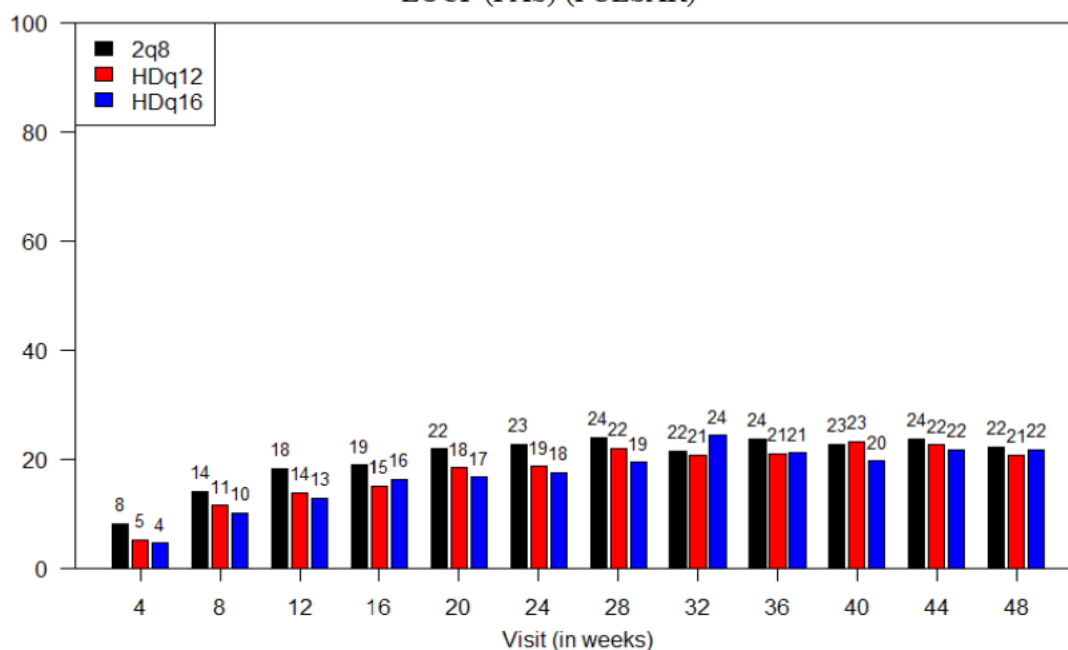
Summary	2q8	HDq12	HDq16	Adjusted Difference (95% CI)	
				HDq12 vs. 2q8	HDq16 vs. 2q8
LOCF (n)	335	334	337		
Gained ≥ 15 letters	74 (22.1)	69 (20.7)	73 (21.7)	-1.8 (-7.8, 4.3)	-0.9 (-7.0, 5.1)
Gained ≥ 10 letters	142 (42.4)	130 (38.9)	130 (38.6)	-3.8 (-11.1, 3.6)	-4.2 (-11.5, 3.0)
Gained ≥ 5 letters	213 (63.6)	186 (55.7)	196 (58.2)	-8.0 (-15.4, -0.7)	-5.5 (-12.9, 1.9)
Gained ≥ 0 letters	267 (79.7)	257 (76.9)	266 (78.9)	-2.7 (-9.0, 3.5)	-0.6 (-6.8, 5.5)
Avoided Losing ≥ 5 letters	301 (89.9)	298 (89.2)	291 (86.4)	-0.6 (-5.2, 4.1)	-3.4 (-8.2, 1.5)
Avoided Losing ≥ 10 letters	317 (94.6)	311 (93.1)	308 (91.4)	-1.5 (-5.1, 2.2)	-3.1 (-6.9, 0.7)

Summary	2q8	HDq12	HDq16	Adjusted Difference (95% CI)	
				HDq12 vs. 2q8	HDq16 vs. 2q8
Avoided Losing ≥ 15 letters	324 (96.7)	318 (95.2)	319 (94.7)	-1.5 (-4.4, 1.5)	-2.0 (-5.1, 1.1)
Observed Cases (n)	285	299	289	--	--
Gained ≥ 15 letters	65 (22.8)	61 (20.4)	62 (21.5)	-2.3 (-8.7, 4.2)	-1.8 (-8.4, 4.7)
Gained ≥ 10 letters	121 (42.5)	118 (39.5)	114 (39.4)	-2.9 (-10.7, 5.0)	-3.4 (-11.3, 4.5)
Gained ≥ 5 letters	184 (64.6)	172 (57.5)	171 (59.2)	-7.0 (-14.8, 0.8)	-5.5 (-13.4, 2.5)
Gained ≥ 0 letters	226 (79.3)	234 (78.3)	231 (79.9)	-1.0 (-7.7, 5.6)	0.8 (-5.8, 7.3)
Avoided Losing ≥ 5 letters	256 (89.8)	271 (90.6)	253 (87.5)	0.8 (-4.0, 5.6)	-2.1 (-7.3, 3.0)
Avoided Losing ≥ 10 letters	269 (94.4)	281 (94.0)	265 (91.7)	-0.4 (-4.2, 3.4)	-2.6 (-6.7, 1.5)
Avoided Losing ≥ 15 letters	276 (96.8)	287 (96.0)	275 (95.2)	-0.9 (-3.9, 2.1)	-1.6 (-4.8, 1.6)

Source: Based on Reviewer's analysis. Abbreviation: LOCF – Last Observation Carried Forward; CMH: Cochran-Mantel-Haenszel test.
Note: Treatment differences and two-sided 95% CIs were based on stratified CMH test using the CMH weights and normal approximation of the weighted estimates. CMH differences and corresponding CIs presented in the table are in percentage.

Figure 5 displays the proportion of subjects who gained ≥ 15 letters from baseline at each visit through Week 48 in the PULSAR study. As shown, the proportion of subjects who gained ≥ 15 letters appeared comparable across the groups at all the visits except that 2q8 treated subjects displayed a small numerical advantage.

Figure 5: Proportion of Subjects Who Gained ≥ 15 Letters from Baseline through Week 48 based on LOCF (FAS) (PULSAR)



Source: Based on Reviewer's analysis.

A supportive analysis conducted using the MI approach led to a similar conclusion (Appendix 1, Figure 19, Figure 20, and Table 40).

Secondary Efficacy Endpoint: Change from baseline in Central Subfield Retinal Thickness (CST) at Week 48

The change in central subfield retinal thickness (CST) from baseline at Week 48 was assessed as a secondary endpoint. Table 14 displays the least square mean changes (in μm) in CST from baseline at Week 48 and their corresponding 95% confidence intervals, based on observed cases analyzed with MMRM in the full analysis set and based on LOCF approach analyzed using ANCOVA. As shown, it appeared that HDq12 and HDq16 treated subjects displayed numerically better reductions at Week 48 from baseline compared to 2q8 treated subjects.

Table 14: Least Square Mean Change in CST from Baseline at Week 48, FAS (PULSAR)

		2q8	HDq12	HDq16
Main Analysis: Using MMRM method				
OC in FAS	LS Mean (95% CI)	-136.3 (-144.6, -127.9)	-147.4 (-155.3, -139.5)	-146.8 (-154.2, -139.4)
	Diff (95% CI)	---	-11.1 (-21.1, -1.2)	-10.5 (-20.1, -0.9)
Supporting Analysis: Using ANCOVA method				
LOCF in FAS	LS Mean (95% CI)	-139.6 (-148.1, -131.2)	-147.4 (-155.9, -138.9)	-148.5 (-156.9, -140.0)
	Diff (95% CI)	---	-7.8 (-17.2, -1.6)	-8.8 (-18.2, 0.5)

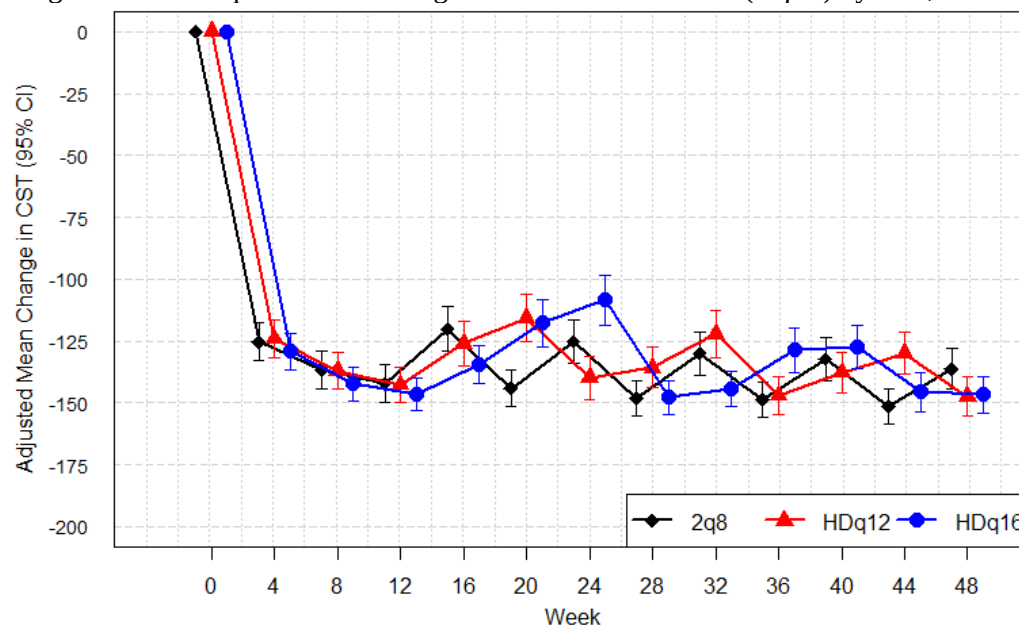
OC: Observed cases; LOCF: Last observation carried forward; MMRM: Mixed model repeated measurements; FAS: Full analysis set.

Source: Reviewer's analysis based on ADEFF.xpt dataset located at

[\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

Figure 6 displays the least squares mean changes in CST over time, based on the observed cases analyzed with MMRM. As shown, the adjusted means decreased from baseline in CST through visits, in similar trajectories among treatment groups.

Figure 6: Least Square Mean Changes from Baseline in CST (in μm) by visit, MMRM, FAS (PULSAR)



Source: Based on Reviewer's analysis.

Reviewer's Comment:

Figure 6 above sheds some light on the impact of frequent dosing versus less frequent dosing on the CST outcome over time. As shown, up to Week 12 where all groups received fixed monthly dosing, a comparable and smooth continued reduction in CST was observed in all groups. However, a wave-like pattern and a shift in these patterns observed beyond Week 12 in each treatment group was likely attributed to the less frequent dosing. For example, beyond Week 12, the reduction in CST worsened up until the next injection visit due to extended dosing (e.g., at Week 16 for 2q8, at Week 20 for HDq12, and at Week 24 for HDq16) and then a better reduction was observed in the next immediate visit after the injection (e.g., at Week 20 in 2q8, at Week 24 in HDq12, and at Week 28 in HDq16). The same pattern continued through Week 48.

3.2.1.5. Efficacy Conclusion

Based on the efficacy evidence from the PULSAR study, the reviewer concludes that subjects treated with aflibercept 8 mg administered every 12 weeks (HDq12) and every 16 weeks (HDq16) after three initial monthly doses had a noninferior mean change in BCVA from baseline at Week 48 compared to subjects treated with aflibercept 2 mg administered every 8 weeks after three initial monthly dosing (2q8). It should be noted that although both HDq12 and HDq16 were noninferior to 2q8 in the primary endpoint, subjects treated with 2q8 displayed numerically better BCVA gain at each visit through Week 48 compared to subjects treated with HDq12 and HDq16 (*Figure 3*).

Additionally, in the PULSAR study, the high dose of aflibercept (pooled HDq12 and HDq16) were superior to 2q8 in the key secondary endpoint of the proportion of subjects with no IRF and no SRF at Week 16. Although superiority in this endpoint was established at Week 16, consistent efficacy findings were not achieved during the study (*Figure 18* and *Table 39* in Appendix 1). For example, at Week 20, the conclusion of superiority reversed and 2q8 treated subjects displayed superior efficacy benefit in this endpoint compared to HDq12 and HDq16 treated subjects. Thus, due to the inconsistent efficacy findings during the study, the treatment benefit of the high doses of aflibercept to 2q8 in the key secondary efficacy endpoint in the PULSAR study may be questionable. The reviewer defers to the medical review team for further input.

3.2.2. PHOTON Study for DME/DR Indications

3.2.2.1. Study Design and Endpoints

Study Design

Data intended to support efficacy and safety for HD aflibercept (8 mg) versus aflibercept 2 mg for the treatment of DME/DR was based on an ongoing phase 2/3, multi-center, randomized, double-masked, active-controlled noninferiority (NI) global study: the PHOTON study.

The primary objective of the PHOTON study was to determine if treatment with HD aflibercept at intervals of 12 or 16 weeks provided noninferior BCVA compared to aflibercept 2 mg dosed every 8 weeks. Of note, aflibercept administered every 8-week interval after five initial monthly injections was approved in the United States on July 2014 for the treatment of DME and on May 2019 for the treatment of DR.

In the PHOTON study, a total of 660 diabetic subjects at least 18 years of age with DME with central involvement who had baseline BCVA letter score of 78 to 24 were enrolled. Eligible subjects were randomized in a 1:2:1 ratio and were to receive aflibercept 2 mg every 8 weeks (2q8, n=167), aflibercept 8 mg every 12 weeks (HDq12, n=329), or aflibercept 8 mg every 16 weeks (HDq16, n=164). Subjects were enrolled in 138 centers in Canada (N = 10), Czech Republic (N = 38), Germany (N = 3), Hungary (N = 77), Japan (N = 74), United Kingdom (N = 4), and the United States (N = 454). Randomization was stratified by baseline CRT (<400 µm vs. ≥400 µm), prior use of DME treatment (yes vs. no), and geographical region (Japan vs. Rest of the World).

The total treatment duration of the PHOTON study is approximately 96 weeks. During this period, subjects assigned in the 2q8 arm were to receive aflibercept 2 mg every 8 weeks following five initial monthly doses, and subjects assigned in the HDq12 and HDq16 arms were to receive aflibercept 8 mg every 12 and every 16 weeks, respectively, following 3 initial monthly doses. For masking purposes, sham injections were given at all visits where an active injection was not planned. During the 96-week treatment period, all subjects were followed every 4 weeks through Week 96.

Although the total study duration of PHOTON study is 96 weeks, this BLA submission was based on the first 48 weeks of efficacy and safety data - the remaining portion of the study is still ongoing. PHOTON was initiated on June 29, 2020, and the data cut-off date for this submission was August 19, 2022.

In the PHOTON study, as in the PULSAR study, dosing intervals for subjects randomized in the HDq12 or HDq16 arms were shortened during the first 48 weeks starting at Week 16 based on protocol-defined dose regimen modifications (DRM) criteria (Appendix 2, [Table 48](#)).

- Subjects in the HDq12 or the HDq16 arm that met the DRM criteria at Week 16 or Week 20 received 8 mg aflibercept at that visit and every 8 weeks thereafter.

- Subjects in the HDq16 arm who did not meet the DRM criteria at Week 16 or 20 but met the criteria at Week 24 received 8 mg aflibercept at that visit and every 12 weeks thereafter.
- Subjects whose interval was not shortened to every 8-week dosing at or before Week 24, the interval was shortened if the DRM criteria was met at a subsequent dosing visit.
 - Subjects in the HDq12 arm who met the DRM criteria received the planned dose at that visit and every 8 weeks thereafter.
 - Subjects in the HDq16 arm who met the DRM criteria received the planned dose at that visit and continued to be dosed every 12 weeks if they were on a 16-week interval or switched to every 8 weeks if they were previously shortened to a 12-week interval.

Figure 7 displays the dosing schedule for PHOTON study.

Figure 7: Dosing schedule for PHOTON study

	Day 1	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24*	Wk 28	Wk 32	Wk 36	Wk 40	Wk 44	Wk 48	1 st Endpoint
2q8	X	X	X	X	X	o	X	o	X	o	X	o	X	
HDq12	X	X	X	o	o [^]	X [^]	o	o	X ^a	o	o	X ^a	o	
HDq16	X	X	X	o	o [*]	o [*]	X [*]	o	o	o	X ^a	o	o	

Dose Regimen Modifications in Year 1 [^] Q12 group: If criteria are met, patients will continue q8. [*] HDq16 group: If criteria are met at week 16 or 20, patient will continue q8. If criteria are met at week 24, patient will continue q12. ^a For patients on a dosing interval of q12 or q16 weeks, DRM criteria will be assessed at dosing visits and if DRM criteria are met the next dosing interval will be reduced by 4 weeks.													
---	--	--	--	--	--	--	--	--	--	--	--	--	--

	Wk 52	Wk 56	Wk 60	Wk 64	Wk 68	Wk 72	Wk 76	Wk 80	Wk 84	Wk 88	Wk 92	Wk 96
2q8	o	X	o	X	o	X	o	X	o	X	o	
HDq12	o	X ^a	o	o	X ^a	o	o	X ^a	o	o	X ^a	
HDq16	o	X ^a	o	o	o	X ^a	o	o	o	X ^a	o	

Dose Regimen Modifications in Year 2 ^a Patients that continue on a dosing interval >8 weeks will be assessed at their dosing visits for DRM criteria for both shortening and extension of the interval by 4 week increments												
--	--	--	--	--	--	--	--	--	--	--	--	--

Abbreviations: 2q8= Aflibercept 2 mg administered every 8 weeks after 5 initial injections at 4-week intervals; HD=high dose; HDq12= High dose aflibercept 8 mg administered every 12 weeks after 3 initial injections at 4-week intervals; HDq16= High dose aflibercept 8 mg administered every 16 weeks after 3 initial injections at 4-week intervals; DRM=Dose regimen modification.

Note: Figure does not reflect all dosing options once a patient is shortened or extended

Source: Figure 2 of PHOTON Clinical Study Report.

Efficacy Evaluation

Key efficacy evaluation in the PHOTON study was based on (i) BCVA letter score as measured using the ETDRS protocol at 4 meters at each study visit, and (ii) diabetic retinopathy severity as measured by the diabetic retinopathy severity score (DRSS).

The DRSS, a validated method measuring changes in DR, characterizes retinopathy based on assessment of abnormalities in seven defined fields of fundus photographs (FP). It is graded according to a 10-step severity score dividing DR severity into levels ranging from DR absent to high-risk proliferative diabetic retinopathy (PDR) (see [Table 18](#) for detail).

In the PHOTON study, BCVA was measured every 4 weeks through Week 48 in the study eye and DRSS was measured at baseline and Weeks 12, 24, 36, and 48 in the study eye.

Study Endpoints

The primary and key secondary efficacy endpoints were *the change from baseline in BCVA at Week 48* and *the proportion of subjects with a ≥ 2 -step improvement in DRSS at Week 48*, respectively.

Additional secondary efficacy endpoints included:

- Proportion of subjects gaining ≥ 15 letters in BCVA at Week 48.
- Proportion of subjects with BCVA ≥ 69 letters at Week 48.
- Proportion of subjects without fluid at foveal center at Week 48.
- Change from baseline in central retinal thickness (CRT) at Week 48.
- Proportion of subjects without leakage on fluorescein angiography at Week 48.
- Change from baseline in National Eye Institute Visual Function Questionnaire (NEI-VFQ) total score at Week 48.

3.2.2.2. Statistical Methodologies

The primary analysis was based on the estimand concept. The primary estimand of interest was based on a combination of hypothetical and treatment policy strategies. It described the change from baseline for all subjects who started treatment, under a pre-specified data handling strategies for occurrence of intercurrent events (see detail below).

Sample Size Determination

The Applicant planned to enroll a total of 640 subjects (160, 160, and 320 subjects in the 2q8, HDq16, and HDq12 arms, respectively). This sample size provided, under the hierarchical testing procedure, at least 98% power for comparing HDq12 vs. 2q8 and subsequently, at least 92% power for comparing HDq16 vs. 2q8, using a noninferiority margin of 4 letters.

The sample size calculation was based on a 1-sided t-test assuming a true treatment difference of zero and a common standard deviation of 9.07 letters for the mean change in BCVA from baseline at Week 48, with one-sided significance level $\alpha=0.025$ and a 19% dropout rate.

Analysis Populations

Three analysis populations were defined for the analyses of the efficacy and safety variables in PHOTON study: (i) full analysis set (FAS) - included all randomized subjects who received at least 1 dose of study medication; it is based on the treatment assigned to the patient at baseline (as randomized), (ii) per-protocol set (PPS) - included all subjects in the FAS who had a baseline and at least 1 post-baseline assessment of BCVA, and did not have any relevant important protocol violations that affected the primary efficacy variable, and (iii) safety analysis set (SAF) - included all randomized subjects who received any study treatment; it is based on the treatment received (as treated).

The primary analysis set for efficacy assessment was the FAS. Supplementary analysis based on the PPS were considered as supporting.

Justification of the Noninferiority Margin

The primary efficacy objective of PHOTON study was to demonstrate noninferiority of HDq16 and HDq12 to 2q8 in the mean change in BCVA from baseline at Week 48 using a noninferiority margin of -4.0 letters. And the secondary efficacy objective of the study was to demonstrate noninferiority of HDq16 and HDq12 to 2q8 in the key secondary efficacy variable of the proportion of subjects who achieved ≥ 2 step-improvement in DRSS from baseline at Week 48 using a noninferiority margin of -15%.

In the study protocol and corresponding statistical analysis plan, the Applicant provided a reasonable justification for the NI margin of -4.0 letters proposed for the primary efficacy variable. However, during the PHOTON protocol review, the Agency indicated that a noninferiority margin of -15% for the key secondary efficacy variable was not acceptable.

Reviewer's Comment:

Based on the efficacy data from previous studies with anti-VEGF therapies (e.g., Lucentis [ranibizumab] and Eylea [aflibercept]), the Agency accepts a noninferiority margin of -4.0 letters for the primary efficacy variable of the mean change in BCVA from baseline at Week 48 and a noninferiority margin of -10% for the key secondary efficacy variable of the proportion of subjects who achieved ≥ 2 step-improvement in DRSS from baseline at Week 48.

Primary Estimand

The Applicant's primary estimand of interest was the difference in least squares mean change in BCVA from baseline to Week 48 between each high dose of aflibercept (HDq12 and HDq16) versus aflibercept 2 mg (2q8), for subjects in FAS under the following data handling strategies for those who had occurrence of intercurrent events (ICEs):

- Data for subjects who (i) discontinued study intervention but stayed in the study, (ii) missed an active injection and did not get it at next visit, (iii) used a prohibited medication after baseline visit, or (iv) received a sham injection instead of an active were

censored and assumed that these subjects adhered to the treatment under a hypothetical estimand strategy.

- Data for subjects who discontinued study intervention and study at the same time was also assumed that these subjects adhered to the treatment under a hypothetical strategy.
- Data for subjects who (i) missed an active injection and got it at next visit, (ii) received an active injection instead of sham, or (iii) received a wrong dose level were used regardless of the occurrence of ICE under the treatment policy estimand strategy.

Table 15 displays the summary of subjects with occurrence of ICEs through Week 48 in the PHOTON study.

Table 15: Summary of Intercurrent Events Through Week 48 in PHOTON study

ICE Category	2q8	HDq12	HDq16
Subjects with at least one ICE	14 (8.4%)	48 (14.6%)	13 (8.0%)
Missed injection for any reason before Week 48 (Planned to be active, make-up or scheduled active injection not given at next visit)	14 (8.4%)	47 (14.3%)	13 (8.0%)
Wrong study intervention before Week 48 (Sham injection instead of active)	0	1 (0.3%)	0
Use of a prohibited medication after the first dose of study medication by Week 48, inclusive	0	0	0

Source: Post-text Table 14.2.1.17 of PHOTON Study.

Reviewer's Comment:

In the PHOTON study, few subjects experienced intercurrent events during the 48-week treatment period. Thus, the impact of different data handling strategies on the overall conclusion expected to be very minimal.

Primary Efficacy Analysis for Primary Estimand

The primary efficacy analysis was an evaluation of noninferiority of each high dose of aflibercept (HDq12 and HDq16) to 2q8 in the primary efficacy variable using MMRM method based on the FAS. The model included baseline BCVA measurement as a covariate and treatment group, baseline CRT category, prior DME treatment, geographical region, and visit as fixed effects as well as visit-by-treatment group and visit-by-baseline BCVA interactions. A Kenward-Roger approximation was used for the denominator degrees of freedom. The model assumed an unstructured covariance structure. Only data up to Week 48 was used in this analysis.

Based on the MMRM, either HDq12 or HDq16 dosing regimens was considered noninferior to 2q8 dosing regimen if the lower limit of the two-sided 95% confidence interval (CI) estimates for the treatment difference in the least squares mean change in BCVA from baseline to Week 48 (*high dose minus aflibercept 2 mg*) was greater than -4 letters.

Superiority of each high dose of aflibercept (HDq12 or HDq16) to 2q8 in the primary efficacy variable was also assessed if (i) noninferiority of each high dose of aflibercept to 2q8 in the

primary efficacy variable and (ii) noninferiority of each high dose of aflibercept to 2q8 in the key secondary efficacy variable were established.

Sensitivity Analyses to the Primary Estimand

The Applicant performed the following sensitivity analyses for the primary estimand to assess the robustness of the primary efficacy analysis in the presence of missing data or occurrence of ICE:

- Last observation carried forward (LOCF) was conducted for subjects who had at least one post-baseline value but discontinued prematurely or had occurrence of ICE before Week 48 and ANCOVA was applied for the change from baseline in BCVA at Week 48.
- Multiple imputation (MI) approach assuming a missing at random (MAR) missing data mechanism was implemented and ANCOVA was applied for the change from baseline in BCVA at Week 48 based on the multiply imputed data. Additionally, a tipping point analysis was performed to assess the validity of the MAR assumption.

Secondary Efficacy Analysis: Key Secondary Endpoint

The secondary efficacy analysis was an evaluation of noninferiority of each high dose of aflibercept (HDq12 and HDq16) to 2q8 in the key secondary efficacy variable of the proportion of subjects who achieved ≥ 2 step-improvement in DRSS from baseline at Week 48 based on a stratified Cochran-Mantel-Haenszel (CMH) test, stratified by baseline CRT, prior DME treatment, and geographical region. The Applicant set a margin of -15% for noninferiority evaluation. However, in this review, noninferiority evaluation was made based on the Agency's recommended margin of -10%.

Based on the CMH test, a weighted point and two-sided 95% CI estimates for the difference in proportions (each high dose minus 2q8) using the CMH weights and normal approximation of the weighted estimates was provided, and a given high dose of aflibercept was considered noninferior to 2q8 if the lower limit of the two-sided 95% CI was greater than -10%.

All other binary secondary efficacy variables were analyzed similarly to the key secondary efficacy variable and all continuous secondary variables were analyzed similarly to the primary efficacy variable.

Reviewer's Comment: *The Agency previously communicated that the noninferiority margin of -15% was not acceptable and recommended a margin of -10% for testing the key secondary hypotheses.*

Type I Error Control (Plan for Multiplicity Adjustment)

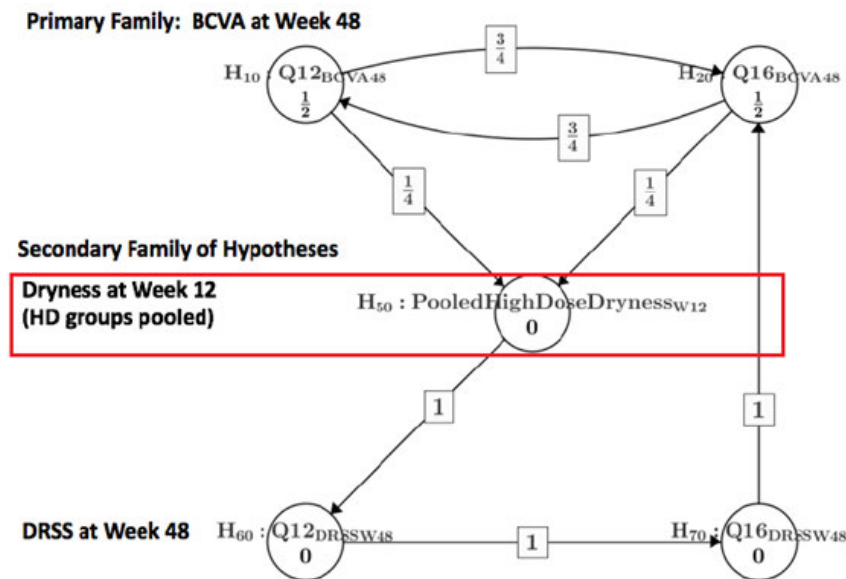
For testing noninferiority of each high dose of aflibercept to 2q8 in the primary and key secondary efficacy endpoints, the Applicant implemented the following hierarchical testing strategy to control the study-wise Type I error rate at two-sided significance level of 5%:

1. Noninferiority of HDq12 to 2q8 in the mean change in BCVA at Week 48.
2. Noninferiority of HDq16 to 2q8 in the mean change in BCVA at Week 48.
3. Noninferiority of HDq12 to 2q8 in the proportion of subjects who achieved ≥ 2 step-improvement in DRSS from baseline at Week 48.
4. Noninferiority of HDq16 to 2q8 in the proportion of subjects who achieved ≥ 2 step-improvement in DRSS from baseline at Week 48.

Reviewer's Comment:

PHOTON study underwent four protocol amendments. Up until protocol amendment 3, the Applicant intended to use the graph-based alpha propagation strategy in Figure 8 to control the study-wise Type I error rate at a two-sided significance level of 5%. As shown, based on the graph-based testing strategy, noninferiority in the DR-related endpoints (denoted by H60 and H70) would only have been tested if noninferiority of either high dose of aflibercept (HDq12 or HDq16) to 2q8 in the primary efficacy endpoint (denoted by H10 and H20) and superiority in dryness endpoint at Week 12 (denoted by H50) were established.

Figure 8: Graphical alpha-propagation approach (PHOTON)



Source: Previous versions of protocol for PHOTON study (up to amendment 3).

In protocol Amendment 4 dated 4/28/22, the Applicant removed the first key secondary endpoint (highlighted in red) in the testing hierarchy **one month before the last patient-last visit date (05/30/2022) and four months before the database lock (8/19/2022)**. The Applicant indicated that the decision to remove the endpoint was based on “recent data from another, completed study”. However, such data and the completed study were not shared with the Agency before. Thus, considering that the removed endpoint showed negative results (Table 29), the Agency questioned the timing of such major protocol change and the reliability of testing the DR-related endpoints in the hierarchy.

To better understand the sequence of events, on 04/03/2023, the Agency submitted the following information request:

In the protocol amendment 4 of the PHOTON study (VGFTe-HD-DME-1934) dated 4/28/2022, you mentioned that the decision of removing the first key secondary endpoint of the “Proportion of patients without retinal fluid at the foveal center at week 12” in the testing hierarchy one month before the last patient-last visit date (05/30/2022) was based on recent data from another, completed study. Regarding such mentioned completed study, provide as many details as possible including when you were aware of the completed study results, who made the decision of removing the endpoint, and whether this person was independent of the ongoing study conduct.

On 04/10/2022, the Applicant provided the following response to the Agency’s request.

“The decision to remove the key secondary endpoint, “Proportion of patients without retinal fluid at the foveal center at week 12” from the PHOTON testing hierarchy was made based on available results from the Phase 2 CANDELA [VGFTe (HD)-AMD-1905] study in patients with nAMD.

... the decision to re-order this specific Week 12 endpoint from the PHOTON study was made internally at Regeneron, after the results from the CANDELA study were fully analyzed and discussed. As the CANDELA and PHOTON studies were run in parallel, the PHOTON study was ongoing when the decision was made and the primary endpoint and initial database lock for PHOTON had not yet been reached. As a result, the study and study teams were still masked to treatment assignment. Although the last patient, last visit was 30 May 2022, the database was not locked and patient data were not unmasked until 19 August 2022.

Throughout the conduct of all trials submitted as the basis of the 8 mg aflibercept BLA, all data capture and masking were maintained at the highest levels in accordance with ICH and FDA guidance. As a result, these trials are adequate and well-controlled and meet all the criteria outlined in CFR 314.126.”

Although the Applicant’s justification appears to be reasonable in this case, such major protocol changes near study end date should be discouraged in general.

Subgroup Analysis

Subgroup analyses of the primary efficacy endpoints were performed by the following subgroup variables: sex, age (<65 years vs. ≥65 years), race (White vs. Others), baseline BCVA category (≤73 vs. >73 letters), prior DME treatment (Yes vs. No), baseline CRT category (<400 vs. ≥400 microns) and region (Japan vs. Rest of the World).

3.2.2.3. *Patient Disposition, Demographic and Baseline Characteristics*

Patient Disposition

Table 16 shows the summary of subject disposition and the primary reasons for study discontinuation during the 48-week treatment period in the PHOTON study. Overall, 660 subjects were randomized. Most subjects completed the 48-week treatment period (> 90%). A total of 47 (7%) subjects discontinued from the study (10 in 2q8 arm, 29 in HDq12, and 8 in HDq16). The discontinuation rate was slightly higher in the HDq12 group compared to in the HDq16 and 2q8 groups. The most common reason for discontinuation was death (2.4%) followed by withdrawal by subject (2.0%) and lost to follow-up (1.1%). In the study, a total of 16 subjects (4 in 2q8, 9 in HDq12, and 3 in HDq16) died during the 48-week treatment period.

Table 16: Summary of Subject Disposition and Reasons for Study Discontinuation (PHOTON)
(All Randomized Subjects)

	2q8	HDq12	HDq16	Total
Randomized, n (%)	167 (100%)	329 (100%)	164 (100%)	660 (100%)
Treated, n (%)	167 (100%)	328 (99.7%)	163 (99.4%)	658 (99.7%)
Completed study at Week 48, n (%)	157 (94.0%)	300 (91.2%)	156 (95.1%)	613 (92.9%)
Discontinued from Study	10 (6.0%)	29 (8.8%)	8 (4.9%)	47 (7.1%)
Primary Reason for Discontinuation				
Adverse Event	0	4 (1.2%)	1 (0.6%)	5 (0.8%)
Death	4 (2.4%)	9 (2.7%)	3 (1.8%)	16 (2.4%)
Lost to Follow-Up	1 (0.6%)	5 (1.5%)	1 (0.6%)	7 (1.1%)
Physician Decision	0	4 (1.2%)	1 (1.2%)	5 (0.8%)
Protocol Deviation	1 (0.6%)	0	0	1 (0.2%)
Withdrawal by Subject	4 (2.4%)	7 (2.1%)	2 (1.2%)	13 (2.0%)

Source: Table 4 of PHOTON Clinical Study Report and Reviewer's analysis.

Analysis Population

Table 17 displays the number of subjects by analysis populations. A total of 660 subjects were randomized in the study. Of these, 2 subjects were removed from the FAS and SAF. A total of 9 subjects in the FAS were excluded from the PPS, and a total of 36 (7 in 2q8, 18 in HDq12, and 11 in HDq16) had major protocol deviations.

Table 17: Summary of Analysis Populations (PHOTON)

	2q8	HDq12	HDq16	Total
Randomized subjects	167	329	164	660
Full Analysis Set (FAS)	167	328	163	658
Safety Analysis Set (SAF)	167	328	163	658
Per-Protocol Set (PPS)	164	322	163	649
Major Protocol Deviations ^(a)	7	18	11	36
Affect primary efficacy variable	3	6	0	9
Does not affect primary efficacy variable	4	12	11	27

(a) Applicant removed from the FAS only those subjects with major protocol deviations that affected the primary efficacy variable.

Source: Table 5 and Table 6 of PHOTON Clinical Study Reports.

Demographic and Baseline Characteristics

The summary of the demographic and baseline characteristics for subjects in the FAS in the PHOTON study are presented in [Table 18](#). Most subjects were male (61%) and white (72%). The average age was about 62 years (range: 24 to 90 years) with a slight majority of subjects (about 56%) were <65 years of age. Most subjects (70%) were from United States and Canada.

The average baseline BCVA in the study eye was 62.5 letters where most subjects had a baseline BCVA of ≤ 73 letters (85%). About 44% of subjects had prior DME treatment in the study eye. Most enrolled subjects in the study had mild-to-severe NPDR (about 80%) and about 10% had mild-to-high risk PDR at baseline.

All demographic and baseline characteristics were well balanced across the three treatment groups.

Table 18: Summary of Demographic and Baseline Characteristics (FAS) (PHOTON)

	2q8 (N = 167)	HDq12 (N = 328)	HDq16 (N = 163)	Total (N = 658)
Age (years)				
Mean (SD)	63.0 (9.8)	62.1 (11.1)	61.9 (9.5)	62.3 (10.4)
Median	64	63	62	63
Range	38 – 90	24 – 87	37 – 83	24 – 90
Age Group				
< 55 years	29 (17.4%)	77 (23.5%)	38 (23.3%)	144 (21.9%)
≥ 55 - < 65 years	63 (37.7%)	108 (32.9%)	54 (33.1%)	225 (34.2%)
≥ 65 - < 75 years	54 (32.3%)	107 (32.6%)	57 (35.0%)	218 (33.1%)
≥ 75	21 (12.6%)	36 (11.0%)	14 (8.6%)	71 (10.8%)
Sex				
Female	75 (44.9%)	118 (36.0%)	64 (39.3%)	257 (39.1%)
Male	92 (55.1%)	210 (64.0%)	99 (60.7%)	401 (60.9%)
Race				
American Indian or Alaska Native	0	2 (0.6%)	0	2 (0.3%)
Asian	30 (18.0%)	48 (14.6%)	23 (14.1%)	101 (15.3%)
Black or African American	18 (10.8%)	35 (10.7%)	9 (5.5%)	62 (9.4%)
Native Hawaiian or Other Pacific Islander	0	1 (0.3%)	0	1 (0.2%)
White	112 (67.1%)	231 (70.4%)	128 (78.5%)	471 (71.6%)
Multiple	0	1 (0.3%)	0	1 (0.2%)
Other	4 (2.4%)	6 (1.8%)	1 (0.6%)	11 (1.7%)
Not Reported	3 (1.8%)	4 (1.2%)	2 (1.2%)	9 (1.4%)
Ethnicity				
Hispanic or Latino	31 (18.6%)	54 (16.5%)	34 (20.9%)	119 (18.1%)
Not Hispanic or Latino	133 (79.6%)	266 (81.1%)	126 (77.3%)	525 (79.8%)
Not Reported	3 (1.8%)	8 (2.4%)	3 (1.8%)	14 (2.1%)
Region				
Japan	20 (12.0%)	37 (11.3%)	17 (10.4%)	74 (11.2%)
Rest of the World	31 (18.6%)	55 (16.8%)	36 (22.0%)	122 (18.5%)
US and Canada	116 (69.5%)	236 (71.9%)	110 (67.5%)	462 (70.2%)

	2q8 (N = 167)	HDq12 (N = 328)	HDq16 (N = 163)	Total (N = 658)
Baseline BCVA (letters)				
Mean (SD)	61.5 (11.2)	63.6 (10.1)	61.4 (11.8)	62.5 (10.9)
Median	63	65	64	65
Range	24 – 78	27 – 79	29 – 78	24 – 79
Baseline BCVA category				
> 73	20 (12.0%)	59 (18.0%)	23 (14.1%)	102 (15.5%)
≤ 73	147 (88.0%)	269 (82.0%)	140 (85.9%)	556 (84.5%)
Baseline CRT (microns)				
Mean (SD)	457.2	449.1	460.3 (117.8)	454.0 (129.5)
Median	417	431	432	430
Range	260 – 1014	229 – 1309	255 – 926	229 – 1309.0
Baseline CRT category				
< 400 microns	72 (43.1%)	134 (40.9%)	65 (39.9%)	271 (41.2%)
≥ 400 microns	95 (56.9%)	194 (59.1%)	98 (60.1%)	387 (58.8%)
Baseline CRT category by IWRS				
< 400 microns	69 (41.3%)	138 (42.1%)	69 (42.3%)	276 (41.9%)
≥ 400 microns	98 (58.7%)	190 (57.9%)	94 (57.7%)	382 (58.1%)
Prior DME treatment in study eye				
Yes	74 (44.3%)	143 (43.6%)	71 (43.6%)	288 (43.8%)
No	93 (55.7%)	185 (56.4%)	92 (56.4%)	370 (56.2%)
Prior DME treatment in study eye by IWRS				
Yes	76 (45.5%)	149 (45.4%)	72 (44.2%)	297 (45.1%)
No	91 (54.5%)	179 (54.6%)	91 (55.8%)	361 (54.9%)
Hemoglobin A1c at baseline category (%)				
≤ 8%	90 (53.9%)	193 (58.8%)	106 (65.0%)	389 (59.1%)
>8%	76 (45.5%)	133 (40.5%)	55 (33.7%)	264 (40.1%)
Missing	1	2	2	5
Diabetes type, n (%)				
Type I	11 (6.6%)	18 (5.5%)	9 (5.5%)	38 (5.8%)
Type II	156 (93.4%)	310 (94.5%)	154 (94.5%)	620 (94.2%)
DR Status (Code)				
1. DR Absent (10, 12)	0	3 (0.9%)	2 (1.2%)	5 (0.8%)
2. DR Questionable (14, 15, 20)	5 (3.0%)	14 (4.3%)	3 (1.8%)	22 (3.3%)
3. Mild NPDR (35)	66 (39.5%)	121 (36.9%)	66 (40.5%)	253 (38.4%)
4. Moderate NPDR (43)	34 (20.4%)	59 (18.0%)	36 (22.1%)	129 (19.6%)
5. Moderately Severe NPDR (47)	17 (10.2%)	46 (14.0%)	15 (9.2%)	78 (11.9%)
6. Severe NPDR (53)	22 (13.2%)	34 (10.4%)	11 (6.7%)	67 (10.2%)
7. Mild NPDR (61)	9 (5.4%)	20 (6.1%)	9 (5.5%)	38 (5.8%)
8. Moderate PDR (65)	4 (2.4%)	11 (3.4%)	9 (5.5%)	24 (3.6%)
9. High Risk PDR (71)	1 (0.6%)	1 (0.3%)	2 (1.2%)	4 (0.6%)
10. High Risk PDR (75)	0	1 (0.3%)	0	1 (0.2%)
90. Cannot Grade	9 (5.4%)	18 (5.5%)	10 (6.1%)	37 (5.6%)

Source: Tables 7 and 8 of PHOTON Clinical Study Report and Reviewer's analysis.

3.2.2.4. Results and Conclusions

Primary Efficacy Endpoint: Change in BCVA from Baseline at Week 48

The primary efficacy objective of the PHOTON study was an assessment of noninferiority of each high dose of aflibercept (HDq12 and HDq16) to aflibercept 2 mg (2q8) in the primary efficacy endpoint in the FAS. The primary estimand of interest to address this objective was the difference in least squares mean change in BCVA from baseline to Week 48 for subjects in FAS under a combination of hypothetical and treatment policy strategy for the handling of intercurrent events.

Table 19 below shows the number of subjects with observed BCVA data in the study at each visit. As stated in Section 3.2.2.2 (RE: Primary Estimand), data under the hypothetical strategy were censored and data under the treatment policy strategy were included in the summary table.

Table 19: Number of Subjects with Observed BCVA Data by Visit (PHOTON)

Treatment	Visit (in Weeks)												
	0	4	8	12	16	20	24	28	32	36	40	44	48
2q8	167	163	163	165	163	147	160	150	159	142	152	141	150
HDq12	328	326	317	314	299	304	272	285	287	271	267	280	277
HDq16	163	163	161	160	159	152	155	144	149	146	151	145	149

Source: Based on Reviewer's analysis.

Based on the observed data in the combination of hypothetical and treatment policy strategy, Table 20 shows the summary of the mean change in BCVA from baseline to Week 48, the adjusted mean estimates for each treatment group at this visit, and the treatment differences in the adjusted means from the MMRM method.

Table 20: Adjusted Mean Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)

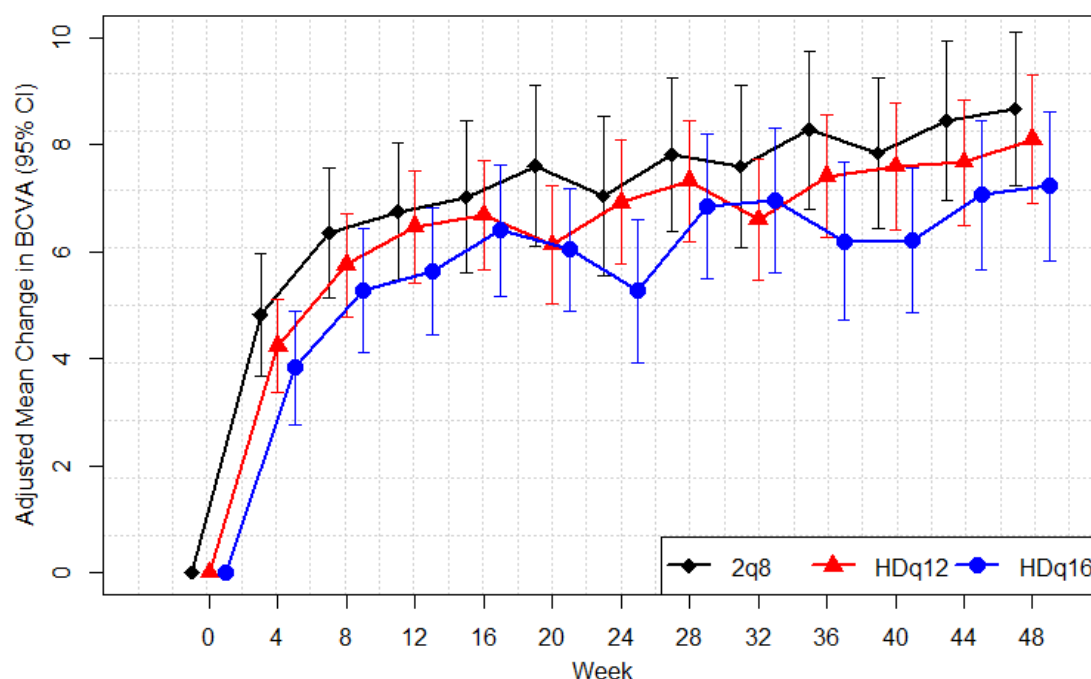
Analysis Visit		2q8 (N = xxx)	HDq12 (N = xxx)	HDq16 (N = xxx)
Mean BCVA				
Baseline	Mean (SD)	61.5 (11.22)	63.6 (10.10)	61.4 (11.76)
	Median (Range)	63 (24 – 78)	65 (27 – 79)	64 (29 – 78)
Week 48	Mean (SD)	71.0 (12.11)	72.6 (11.00)	69.8 (13.20)
	Median (Range)	72.5 (24 - 90)	75 (30 - 95)	72 (16 - 91)
Mean Change in BCVA				
Week 48	Mean (SD)	9.2 (8.99)	8.8 (8.95)	7.9 (8.38)
	Median (Range)	9.5 (-25 – 48)	9 (-33 – 41)	7 (-17 – 40)
Adjusted Mean Change in BCVA				
Week 48	LS Mean (95% CI)	8.7 (7.2, 10.1)	8.1 (6.9, 9.3)	7.2 (5.8, 8.6)
	Diff (95% CI)	---	-0.6 (-2.3, 1.1)	-1.4 (-3.3, 0.4)

Source: Reviewer's analysis based on ADOEVA.xpt dataset located at [\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\vgftehd-dme-1934\analysis\adam\datasets](#) for PHOTON.

As shown in Table 20, subjects treated with high doses of aflibercept (HDq12 or HDq16) had a noninferior mean change in BCVA from baseline at Week 48 compared to subjects treated with aflibercept 2 mg (2q8) because the 95% lower confidence limits for the differences in the adjusted means between each high dose of aflibercept to aflibercept 2 mg was greater than the noninferiority margin of -4.0 letters. The adjusted means for the change in BCVA from baseline to Week 48 was +8.7 letters in 2q8, +8.1 letters in HDq12, and +7.2 letters in HDq16 with a treatment difference of -0.6 (95% CI: -2.3 to 1.1) between HDq12 and 2q8 and -1.4 (95% CI: -3.3 to 0.4) between HDq16 and 2q8.

Figure 9 displays a plot of the adjusted mean change in BCVA from baseline to each visit through Week 48 including the 95% confidence interval estimates (vertical bars at each visit).

Figure 9: Plot of Adjusted Mean change in BCVA from Baseline through Week 48 (FAS) (PHOTON)



Source: Based on Reviewer's analysis (Appendix 2, Table 41)

As shown, the adjusted mean changes from baseline at each visit through Week 48 appear to be numerically higher in the 2q8 group compared to in both of aflibercept high dose groups. However, the confidence intervals at each visit overlapped between the treatment groups implying a lack of treatment difference. The numerically higher adjusted mean changes that subjects in the 2q8 group exhibited compared to subjects in the high dose groups could be attributed to the frequent injections that subjects in the 2q8 arm received through Week 48.

Reviewer's Remark:

See remark for the PULSAR study regarding the wave-like pattern in Figure 9. As in the PULSAR study, because the wave-like pattern after Week 12 was dependent on the timing of the

dose, BCVA assessment at Week 48 alone may benefit one arm over the other depending on when the last injection was received at the prior visits. Thus, to evaluate the impact of this concern, the reviewer compared the treatment groups based on the change in BCVA data from baseline averaged at Week 40, 44, and 48 (Table 21). As shown, using this endpoint, the noninferiority conclusion was unaltered which confirm the robustness of the primary efficacy result at Week 48.

Table 21: Summary of Mean Change in BCVA from Baseline at Week 40, 44, 48, and averaged at Week 40/44/48 (FAS) (PHOTON)

Analysis Visit	2q8	HDq12	HDq16
Week 40	7.8 (0.7)	7.6 (0.6)	6.2 (0.7)
Week 44	8.5 (0.8)	7.7 (0.6)	7.1 (0.7)
Week 48	8.7 (0.7)	8.1 (0.6)	7.2 (0.7)
Week 40/44/48	8.7 (0.6)	8.1 (0.5)	7.2 (0.6)
Diff (95% CI)	--	-0.5 (-2.1, 1.1)	-1.5 (-3.2, 0.2)

Source: Based on Reviewer's analysis.

Sensitivity/Supporting Analyses to the Primary Efficacy Endpoint

Different sensitivity and supporting analyses were performed to assess the robustness of the primary efficacy analysis results based on the MMRM method. ANCOVA method was used to analyze the primary efficacy endpoint by imputing missing data using LOCF approach and using multiple imputation (MI) approach under MAR assumption (Table 22). Additionally, to check the robustness of the MMRM results if the missing data are not MAR, a tipping point analysis was also performed based on the MI analysis and the results are provided in Appendix 2, Table 42.

As shown in Table 22, the results based on the LOCF and the MI missing data handling approaches were consistent with those of the primary analysis results using MMRM in the FAS.

Table 22: Sensitivity and Supportive Analyses: Mean Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)

		2q8	HDq12	HDq16
Sensitivity Analysis: Using ANCOVA method				
LOCF in FAS	LS Mean (95% CI)	8.4 (6.8, 10.1)	7.5 (6.2, 8.8)	6.9 (5.2, 8.6)
	Diff (95% CI)	---	-1.0 (-2.7, 0.8)	-1.5 (-3.5, 0.4)
MI in FAS	LS Mean (95% CI)	8.4 (6.9, 10.1)	7.9 (6.5, 9.1)	6.9 (5.4, 8.6)
	Diff (95% CI)	---	-0.6 (-2.3, 1.1)	-1.5 (-3.4, 0.4)
Supporting Analysis: Using MMRM method				
PPS in FAS	LS Mean (95% CI)	8.7 (7.3, 10.2)	8.1 (6.9, 9.4)	7.2 (5.8, 8.6)
	Diff (95% CI)	---	-0.6 (-2.3, 1.1)	-1.5 (-3.4, 0.3)

Abbreviations: LOCF: Last observation carried forward; MI: Multiple imputation; PPS: Per Protocol Set; MMRM: Mixed model for repeated measures; FAS: Full analysis set.

Source: Reviewer's analysis based on ADOEVA.xpt dataset located at [\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\vgtfthd-dme-1934\analysis\adam\datasets](#) for PHOTON.

Additionally, results of the tipping point analysis in [Table 42](#) showed that the smallest shift parameter (delta) used as penalty for subjects with missing data in the HD groups, for which noninferiority could not be shown anymore, was delta = -10 letters for HDq12 vs. 2q8 and delta = -6 letters for HDq16 vs. 2q8. The "tipping point" values of -10 letters and -6 letters requiring such a large difference in BCVA at Week 48 between HDq16 and HDq12 vs. 2q8, respectively, do not seem plausible implying the primary analysis of noninferiority conclusion based on MMRM method is robust to the departure of the MAR assumption.

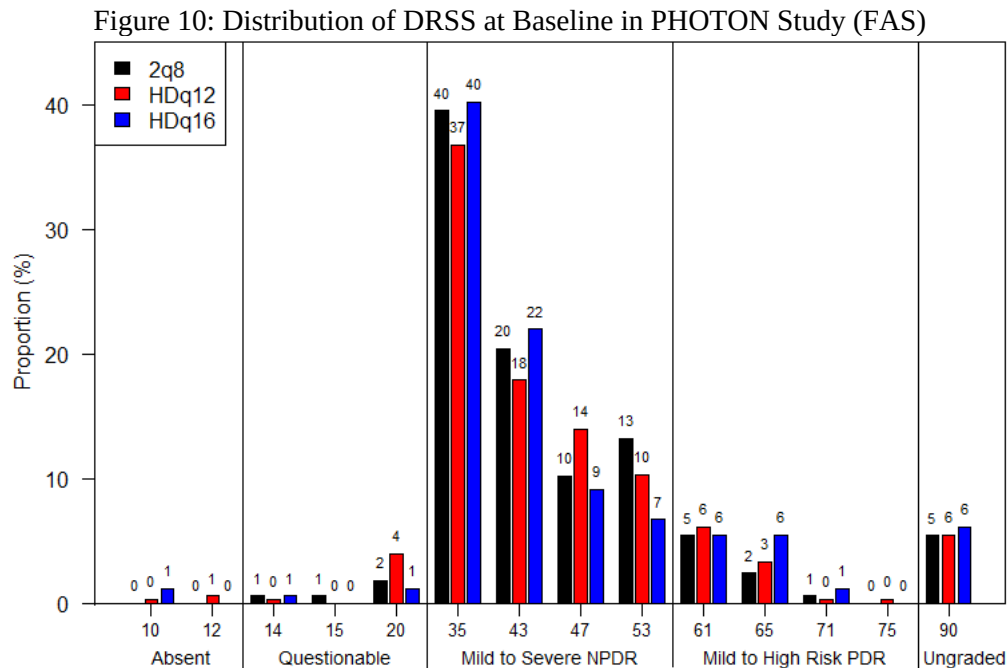
As supporting analysis, analysis of the primary efficacy endpoint was also repeated using the Per Protocol Set (PPS). These results were consistent with the primary efficacy analysis results ([Table 20](#)).

An additional supportive analysis for the primary efficacy endpoint was also conducted based only on subjects who had information at Week 48 and did not have their assigned dosing interval shortened based on the DRM criteria. These results were consistent with the primary efficacy analysis results.

Key Secondary Efficacy Endpoint: Proportion of Subjects who Achieved ≥ 2 -Step Improvement in DRSS from Baseline to Week 48

The proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline to Week 48 was defined as the key secondary efficacy endpoint. This endpoint was intended to support the DR indication.

The distribution of baseline DRSS in each treatment group in the PHOTON study is displayed in [Figure 10](#).



Source: Based on Reviewer's analysis using [Table 18](#).

As shown, most enrolled subjects in the study had mild-to-severe NPDR at baseline (80%) and about 10% had mild-to-high risk PDR. A total of 37 (6%) subjects had ungradable baseline DRSS (9 in 2q8, 18 in HDq12, and 10 in HDq16) and these subjects were not included in the DR-related efficacy analysis. The distribution of baseline DRSS appeared comparable across the treatment groups.

Treatment Comparison in the Key Secondary Endpoint

For the key secondary efficacy endpoint of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline to Week 48, noninferiority of each high dose of aflibercept (HDq12 and HDq16) to aflibercept 2q8 was assessed using a CMH test stratified by baseline CRT category ($<400 \mu\text{m}$, $\geq 400 \mu\text{m}$), prior DME treatment (yes vs. no), and geographical region (Rest of world vs Japan). Similar analysis was performed for the proportion of subjects who achieved ≥ 3 -step improvement in DRSS from baseline to Week 48 as an exploratory endpoint.

In the Applicant analysis of the key secondary endpoint, missing or ungradable postbaseline data were imputed using LOCF procedure. Subjects were considered as non-responders if all post-baseline measurements were missing or ungradable. Strategies for occurrence of intercurrent events were handled similarly to in the primary efficacy endpoint, except that censored data were imputed by LOCF procedure.

Table 23 displays the number of subjects in each treatment group with observed, ungradable, censored, and missing DRSS data at each visit where DRSS data were collected through Week 48. As shown, a total of 621 subjects (158 in 2q8, 310 in HDq12, and 153 in HDq16) had gradable DRSS data at baseline whereas 37 subjects (9 in 2q8, 18 in HDq12, and 10 in HDq16) had ungradable DRSS data at baseline. Accordingly, in this review, all DR-related endpoints were assessed based on subjects with gradable DRSS data at baseline. That is, the 37 subjects with ungradable DRSS data at baseline were excluded in the analyses of all DR-related endpoints.

Table 23: Number of Subjects with Observed DRSS Data from Baseline through Week 48

DRSS Status	Treatment	Baseline	Week 12	Week 24	Week 36	Week 48
Observed	2q8	158	149	159	161	162
	HDq12	310	280	299	308	308
	HDq16	153	143	153	155	155
Ungradable	2q8	9	9	3	0	0
	HDq12	18	23	6	4	3
	HDq16	10	14	6	5	5
Censor	2q8	0	0	0	1	1
	HDq12	0	4	1	0	1
	HDq16	0	0	1	1	1
Missing	2q8	0	3	4	4	4
	HDq12	0	11	16	16	16
	HDq16	0	3	2	2	2

Source: Based on Reviewer's analysis

Table 24 shows the key secondary efficacy analysis results using the LOCF method to impute missing data. As shown, 27%, 29%, and 20% of subjects in the 2q8, HDq12, and HDq16 treatment groups, respectively, achieved ≥ 2 -step improvement in DRSS from baseline to Week 48. Treatment differences for the adjusted proportions was 2.0% (95% CI: -6.4%, 10.4%) for HDq12 versus 2q8 and -7.5% (95% CI: -16.8%, 1.8%) for HDq16 versus 2q8.

HDq12 was noninferior to 2q8 with respect to this endpoint as the lower bound of the 95% CI for the difference in the adjusted proportions between HDq12 and 2q8 of -6.4% was $> -10\%$. However, this could not be shown for the HDq16 group since the lower bound of the 95% CI for the difference in the adjusted proportions between HDq16 and 2q8 of -16.8% was $< -10\%$.

Table 24: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline to Week 48 (Full Analysis Set)

Summary	2q8 (N = 158)	HDq12 (N = 310)	HDq16 (N = 153)
Achieved ≥ 2-step improvement	42 (26.6%)	90 (29.0%)	30 (19.6%)
Adjusted Difference (95% CI)	---	2.0% (-6.4%, 10.4%)	-7.5% (-16.8%, 1.8%)
Achieved ≥ 3-step improvement	23 (14.6%)	37 (11.9%)	14 (9.2%)
Adjusted Difference (95% CI)	---	-2.8% (-9.3%, 3.7%)	-5.6% (-12.8%, 1.6%)

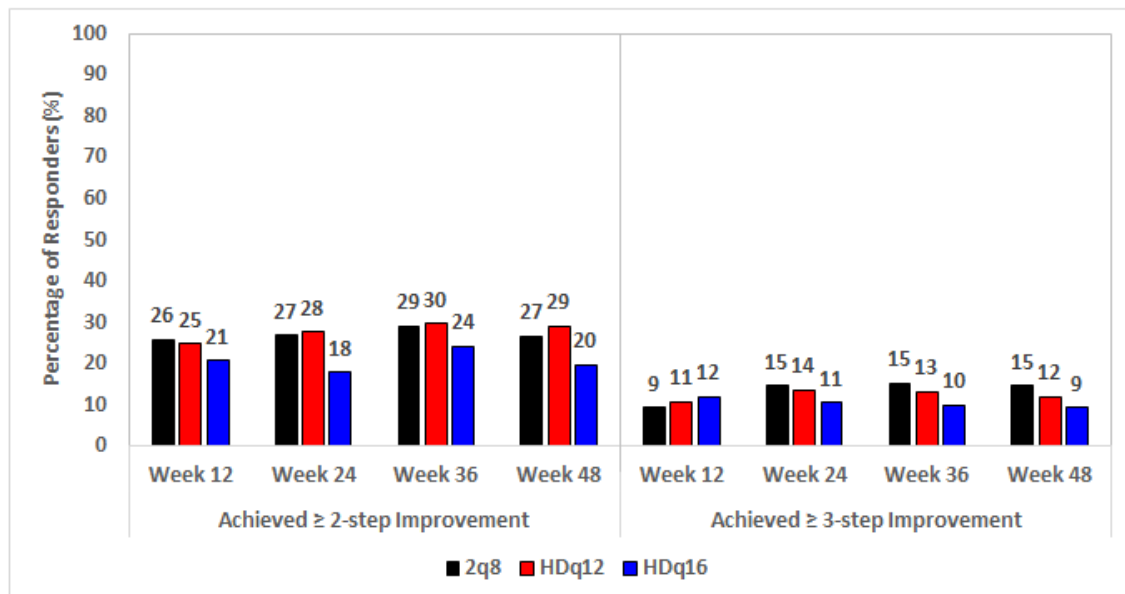
Source: Based on Reviewer's analysis.

LOCF: Last Observation Carried Forward; CMH: Cochran-Mantel-Haenszel test.

Note: Treatment differences and two-sided 97.5% CI were based on stratified CMH test using the CMH weights and normal approximation of the weighted estimates.

Figure 11 displays the proportion of subjects who achieved ≥ 2 -step and ≥ 3 -step improvement in DRSS from baseline at each visit through Week 48.

Figure 11: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline at each Visit (LOCF) (FAS)



Source: Based on Reviewer's analysis.

As shown, subjects in the 2q8 and HDq12 groups displayed comparable improvement at each visit. However, the improvement in the HDq16 group at each visit was numerically lower than those in the 2q8 or HDq12 groups.

Figure 22 in Appendix 2 displays the cumulative distribution of the change in DRSS from baseline at Week 48 and the rate of DR improvement at different cut-off points were displayed in the Figure (*RE: dashed lines*). As shown, a comparable number of subjects in 2q8 and HDq12 achieved improvement in DRSS from baseline at Week 48.

Sensitivity analyses performed using multiple imputation (Appendix 2, Table 44) and observed cases (Appendix 2, Table 45) were consistent with the key secondary efficacy results based on the LOCF approach, leading to the same conclusion for a robust interpretation of the noninferiority finding of HDq12 to 2q8 and the lack of noninferiority finding of HDq16 to 2q8.

Subgroup Analysis of the Key Secondary Endpoint by Baseline DRSS Category

Subgroup analysis of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline at Week 48 was performed by baseline DRSS categories (Figure 17). The baseline DRSS categories included: (i) subjects with moderate NPDR or better (DRSS ≤ 43), (ii) subjects with moderately severe or severe NPDR (DRSS = 47 or 53), and (iii) subjects with at least mild PDR (DRSS ≥ 61).

The distribution of subjects in these baseline DRSS categories are shown in Table 25 below by treatment group. As shown, most subjects in each treatment group had moderate NPDR or better DRSS at baseline (DRSS ≤ 43) but slightly more subjects in the HDq16 group (70%) had DRSS ≤ 43 at baseline than subjects in the 2q8 (67%) and HDq12 (64%) groups.

Table 25: Number of Subjects by Baseline DRSS category and by treatment (FAS) (PHOTON)

Baseline DRSS	2q8 (N = 158)	HDq12 (N = 310)	HDq16 (N = 153)
≤ 43	105 (66.5%)	197 (63.5%)	107 (70.0%)
47 and 53	39 (24.5%)	80 (25.8%)	26 (17.0%)
≥ 61	14 (8.9%)	33 (10.6)	20 (13.1%)

Source: Based on Reviewer's analysis.

In the PHOTON study, the extent of ≥ 2 -step improvement in DRSS from baseline at Week 48 varied by baseline DR severity level (Table 26). As shown, since subjects with baseline DRSS ≤ 43 generally would not have had the same opportunity to achieve ≥ 2 -step improvement in DRSS score as subjects with more severe disease (because there is less room to improve), the improvement in this category was very small and comparable across the treatment groups. On the other hand, in subjects with moderately severe or severe NPDR at baseline (DRSS = 47 or 53), more than 75% of subjects in each treatment group achieved ≥ 2 -step improvement in DRSS from baseline at Week 48 (82% in 2q8, 83% in HDq12, and 77% in HDq16). Subjects who had a severe form of DR at baseline (i.e., DRSS ≥ 61) appeared to have benefitted more from treatment with 2q8 (64%) than with treatment with HDq12 (49%) or HDq16 (45%).

Table 26: Proportion of Subjects who Achieved ≥ 2 -Step Improvement in DRSS from Baseline to Week 48 by Baseline DRSS Category (FAS) (PHOTON)

		2q8 (N = 158)	HDq12 (N = 310)	HDq16 (N = 153)
DRSS ≤ 43	% (n/N)	1.0% (1/105)	4.1% (8/197)	0.9% (1/107)
	Difference (95% CI)	--	3.1% (-0.7%, 6.9%)	-0.0% (-3.0%, 3.0%)
DRSS 47 and 53	% (n/N)	82.1% (32/ 39)	82.5% (66/80)	76.9% (20/26)
	Difference (95% CI)	--	0.4% (-16.3%, 17.2%)	-5.1% (-28.2%, 18.0%)
DRSS ≥ 61	% (n/N)	64.3% (9/ 14)	48.5% (16/33)	45.0% (9/20)
	Difference (95% CI)	--	-16 (-50.5%, 18.9%)	-19% (-57.3%, 18.8%)

Source: Based on Reviewer's analysis.

Reviewer's Comment:

Although HDq16 was not noninferior to the active control 2q8 in the PHOTON study, the magnitude of improvement observed with Q16W in the study was much higher than what has been observed with either sham (15%) or laser treatment (7% - 24%) in previous studies conducted in subjects with DR (Table 27). Additionally, in subjects with severe or severe NPDR at baseline (DRSS = 47 or 53), HDq16 treated subjects in the PHOTON study has comparable ≥ 2 -step improvement (77%) as 2q8 treated subjects in the PANORAMA study (80%).

Table 27: Proportion of subjects Who Achieved ≥ 2 -Step Improvement from Baseline in the DRSS at Year 1

	Overall Population			Subjects with Baseline DRSS = 47 or 53		
	PHOTON	VISTA	VIVID	PHOTON	Pooled VISTA/VIVID	PANORAMA
Sham	--	--	--	--	--	15%
Laser	--	8%	14%	--	16%	--
2q8	27%	29%	28%	82%	50%	80%
HDq12	29%	--	--	83%	--	--
HDq16	20%	--	--	77%	--	--

Note 1: VISTA and VIVID were pivotal Phase 3 studies that provided efficacy and safety support for the approval of aflibercept 2q8 for the treatment of DME and for the treatment DR in patients with DME. The studies enrolled subjects with DME.

Note 2: PANORAMA was a Pivotal Phase 3 study used to provide support for aflibercept 2q8 for the broader indication of treatment of DR regardless of DME. The study enrolled patients with moderate to severe NPDR at baseline without DME.

Secondary Efficacy Endpoint: Proportion of Subjects Who Gained/Avoided Losing Letters in BCVA from Baseline at Week 48

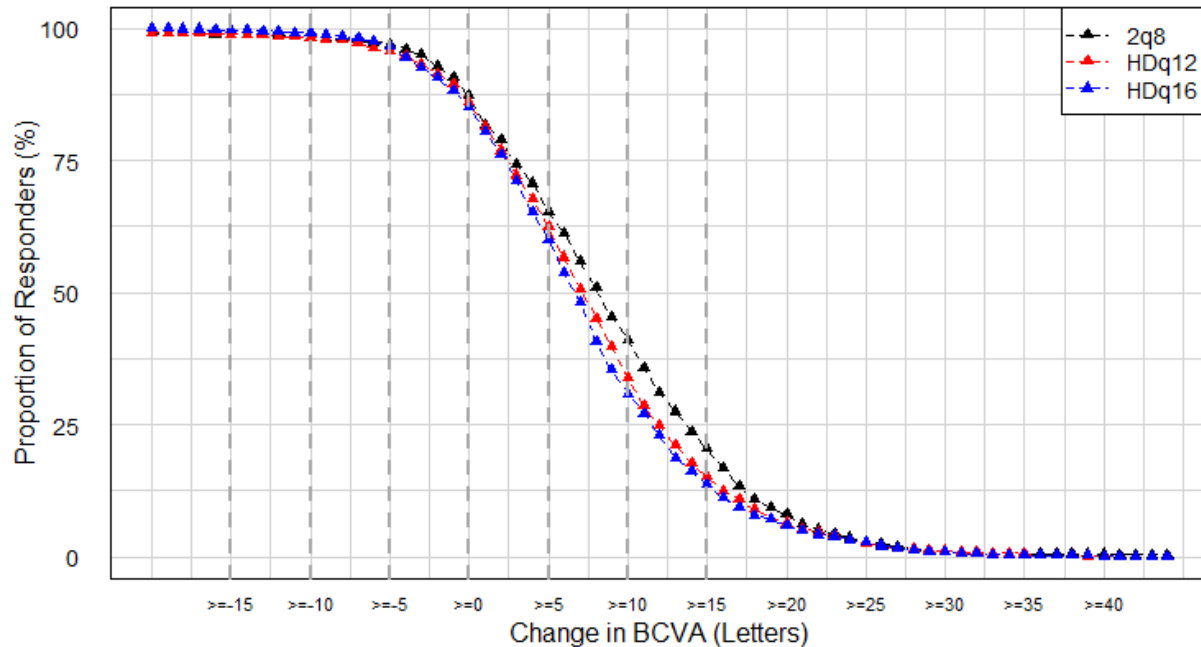
The proportion of subjects who gained ≥ 15 letters in BCVA from baseline at Week 48 was a secondary efficacy endpoint in the PHOTON study. Additionally, the proportion of subjects who gained/avoided losing ≥ 5 or ≥ 10 letters from baseline at Week 48 were defined as exploratory endpoints.

Figure 12 displays the cumulative distribution of the change in BCVA at Week 48 based on missing data imputed using multiple imputation. As shown, the three treatment groups seemed comparable across the binary secondary endpoints (dashed lines).

Table 28 shows the summary of the proportion of subjects who gained/avoided losing letters in BCVA at Week 48 based on missing data imputed using LOCF approach and observed cases. The HD treatment groups appeared to be comparable in all considered endpoints. For example,

using the LOCF approach, 19% of the subjects in the HDq12 group and 17% of the subjects in the HDq16 group gained ≥ 15 letters from baseline at Week 48, compared to 23% in the 2q8 group with treatment differences of -4.6% (95% CI: -12.2% to 2.9%) and -7.1% (-15.2% to 0.9%), respectively. The results based on observed data provided a similar conclusion with minor numerical differences. The same can be determined in the proportion of subjects who avoided losing letters in BCVA. For example, 99% of the subjects in all the groups avoided losing ≥ 15 letters, with treatment differences of -0.3% (95% CI: -2.5% to 1.9%) and -0.7% (95% CI: -1.4%, 2.7%), respectively.

Figure 12: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (FAS) (PHOTON)



Source: Based on Reviewer's analysis.

Table 28: Proportion of Subjects Who Gained ≥ 15 , ≥ 10 , ≥ 5 and ≥ 0 Letters and Avoided Losing ≥ 5 , ≥ 10 and ≥ 15 Letters in BCVA from Baseline at Week 48, FAS (PHOTON)

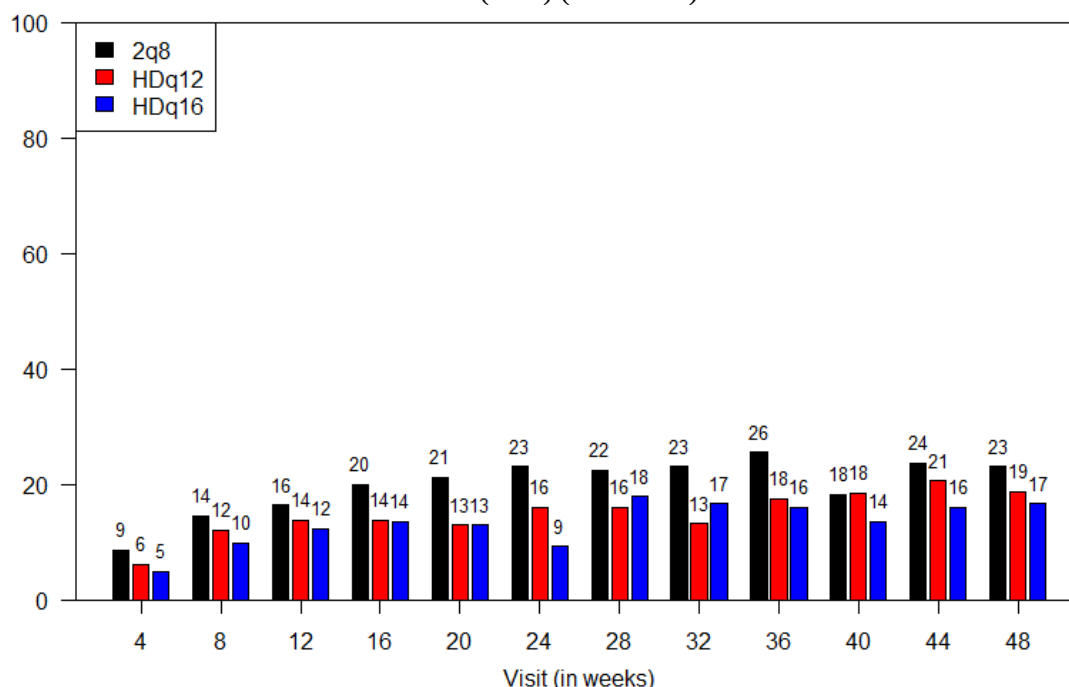
Summary	2q8	HDq12	HDq16	Adjusted Difference (95% CI)	
				HDq12 vs. 2q8	HDq16 vs. 2q8
LOCF (n)	165	236	163	--	--
Gained ≥ 15 letters	38 (23.0)	61 (18.7)	27 (16.6)	-4.6 (-12.2, 2.9)	-7.1 (-15.2, 0.9)
Gained ≥ 10 letters	81 (49.1)	132 (40.5)	57 (35.0)	-9.2 (-18.2, -0.1)	-15.0 (-25.0, -5.1)
Gained ≥ 5 letters	113 (68.5)	231 (70.9)	107 (65.6)	2.0 (-6.4, 10.3)	-3.7 (-13.5, 6.1)
Gained ≥ 0 letters	148 (89.7)	283 (86.8)	144 (88.3)	-3.1 (-8.8, 2.6)	-1.9 (-8.4, 4.5)
Avoided Losing ≥ 5 letters	162 (98.2)	306 (93.9)	156 (95.7)	-4.3 (-7.6, -1.0)	-2.3 (-6.0, 1.3)
Avoided Losing ≥ 10 letters	163 (98.8)	316 (96.9)	161 (98.8)	-1.8 (-4.3, 0.7)	0.1 (-2.3, 2.5)
Avoided Losing ≥ 15 letters	163 (98.8)	321 (98.5)	162 (99.4)	-0.3 (-2.5, 1.9)	0.7 (-1.4, 2.7)
Observed Cases (n)	150	277	149	--	--
Gained ≥ 15 letters	36 (24.0)	55 (19.9)	26 (17.4)	-4.2 (-12.3, 3.9)	-6.8 (-15.4, 1.7)
Gained ≥ 10 letters	75 (50.0)	120 (43.3)	55 (36.9)	-6.8 (-16.4, 2.8)	-13.7 (-24.0, -3.3)
Gained ≥ 5 letters	104 (69.3)	204 (73.6)	99 (66.4)	4.1 (-4.5, 12.8)	-3.3 (-13.4, 6.9)

Summary	2q8	HDq12	HDq16	Adjusted Difference (95% CI)	
				HDq12 vs. 2q8	HDq16 vs. 2q8
Gained ≥ 0 letters	134 (89.3)	244 (88.1)	131 (87.9)	-1.4 (-7.3, 4.6)	-2.0 (-8.8, 4.8)
Avoided Losing ≥ 5 letters	147 (98.0)	264 (95.3)	142 (95.3)	-2.6 (-5.9, 0.7)	-2.6 (-6.6, 1.4)
Avoided Losing ≥ 10 letters	148 (98.7)	273 (98.6)	147 (98.7)	-0.0 (-2.3, 2.3)	0.0 (-2.5, 2.6)
Avoided Losing ≥ 15 letters	148 (98.7)	276 (99.6)	148 (99.3)	1.1 (-0.9, 3.0)	0.7 (-1.5, 2.9)

Source: Based on Reviewer's analysis.

Figure 13 displays the proportion of subjects who gained ≥ 15 letters from baseline at each visit through Week 48 in the PHOTON study. As shown, the proportion of subjects who gained ≥ 15 letters appeared comparable across the groups at most the visits except that 2q8 treated subjects displayed a numerical advantage.

Figure 13: Proportion of Subjects Who Gained ≥ 15 Letters from Baseline through Week 48 based on LOCF (FAS) (PHOTON)



Source: Based on Reviewer's analysis.

A supportive analysis conducted using the MI approach led to a similar conclusion (Appendix 2, Figure 21 and Table 43).

Secondary Efficacy Endpoint: Proportion of Subjects Without Fluid at Foveal Center at Week 48

The proportion of subjects without fluid at foveal center at Week 48 was defined as additional secondary efficacy endpoint (Table 29). As shown, a comparable proportion of subjects in the HDq12 and 2q8 groups were without fluid at foveal center at Week 48. For examples, the proportion of subjects with no fluid at foveal center at Week 48 was 58.5% and 43.8% in the

HDq12 and HDq16 groups, respectively, compared to 54.5% in the 2q8 group with treatment differences of 4.3% (95% CI: -4.7%, 13.3%) and -9.7% (95% CI: -20.3%, 0.8%), respectively.

Table 29: Proportion of Subjects with no Fluid at Foveal Center at Week 48 and Week 12, LOCF (FAS) (PHOTON)

	2q8	HDq12	HDq16
Subjects who had no fluid, n (%)			
Week 48, n/N (%)	90/165 (54.5%)	190/325 (58.5%)	71/162 (43.8%)
Difference % (95% CI)	--	4.3 (-4.7, 13.3)	-9.7 (-20.3, 0.8)
		Pooled (HDq12 and HDq16)	
Week 12, n/N (%)	66/164 (40.2%)	182/486 (37.4%)	
Difference % (95% CI)	--	-2.3 (-10.8, 6.2)	

LOCF: Last observation carried forward; Num/Den: Numerator/Denominator.

Source: Reviewer's analysis based on ADMOOCT.xpt dataset located at [\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\vgftehd-dmc-1934\analysis\adam\datasets](#) for PHOTON.

A sensitivity analysis performed based on the observed data resulted in a similar conclusion as the one with the LOCF approach except for a greater numeric difference in favor of HDq12 (Appendix 2, [Table 46](#)).

[Table 29](#) also presents results of the proportion of subjects with no fluid at foveal center at Week 12 for the pooled aflibercept HD groups (HDq12 and HDq16) and 2q8. For this assessment, HDq12 and HDq16 arms were pooled since at Week 12 both groups were treated identically with three initial monthly doses. As shown, superiority for this endpoint was not established at Week 12.

Because of the negative result for this endpoint, the DR-related key secondary endpoint to support the DR indication would not have been tested per the initial graphical testing procedure described in the original protocol and in protocol amendments 1-3 (see [Section 3.2.2.2 \[RE: Type I Error Control \(Plan for Multiplicity Adjustment\)\]](#) for detail). It should be noted that the endpoint of the proportion of subjects with no fluid at foveal center at Week 12 was defined as a key secondary endpoint up until its removal in protocol amendment 4; that is, one month before the last patient-last visit date and four months before the database lock.

Reviewer's Comment:

Although at Week 48 the individual HDq12 arm showed higher proportion of subjects with no fluid at foveal center compared to 2q8, this results were not observed at each visit through Week 48 ([Figure 23](#) and [Table 47](#) in Appendix 2). For example, more subjects in the 2q8 group (50%) than in the HDq12 group (36%) and in the HDq16 group (35%) had no fluid at foveal center at Week 20 with a treatment difference of -14% (95% CI: -23% to -4%) favoring 2q8 compared to HDq12 and -15% (95% CI: -25% to -4%) favoring 2q8 compared to HDq16.

Secondary Efficacy Endpoint: Change from baseline in Central Retinal Thickness (CST) at Week 48

The change in central subfield retinal thickness (CST) from baseline at Week 48 was assessed as a secondary endpoint. [Table 30](#) displays the least square mean changes (in μm) in CST from

baseline at Week 48 and their corresponding 95% confidence intervals, based on observed cases analyzed with MMRM method in the full analysis set and based on LOCF approach analyzed using ANCOVA. It appeared that HDq12 treated subjects displayed numerically better reductions at Week 48 from baseline than 2q8 treated subjects. The same did not occur to HDq16 treated subjects.

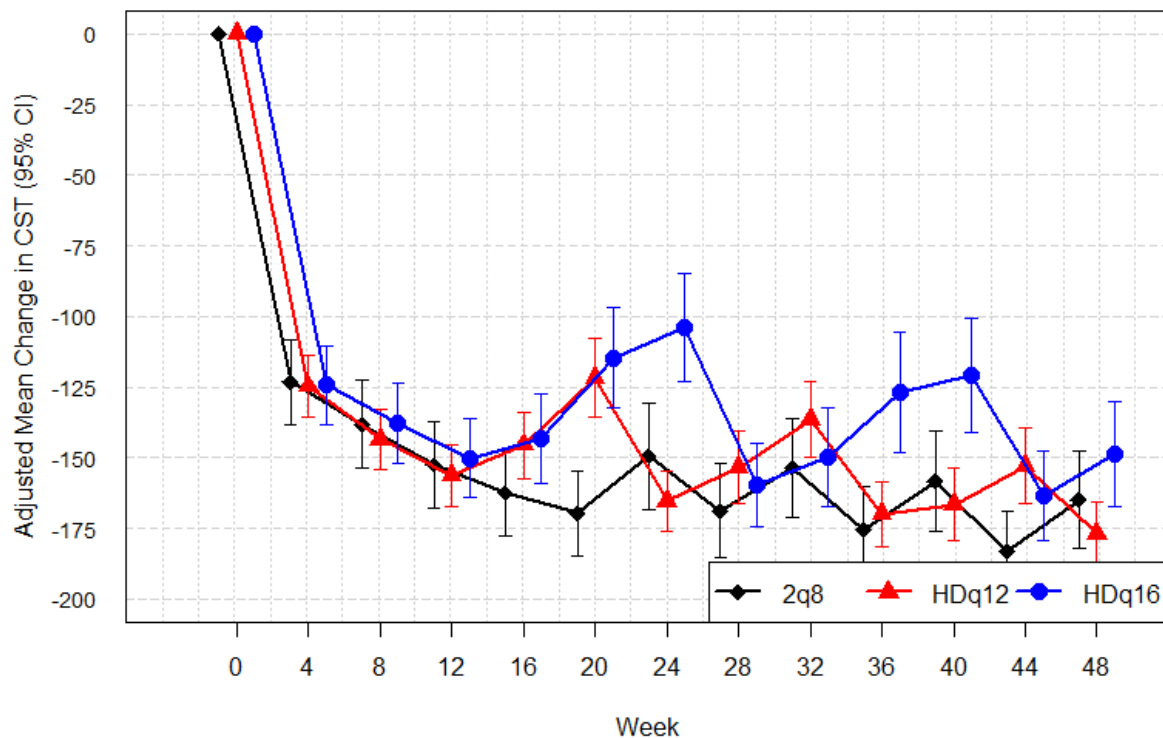
Table 30: Least Square Mean Change in CST from Baseline at Week 48, FAS (PHOTON)

		2q8	HDq12	HDq16
Main Analysis: Using MMRM method				
OC in FAS	LS Mean (95% CI)	-164.8 (-182.2, -147.5)	-176.8 (-188.0, -165.5)	-148.8 (-167.5, -130.2)
	Diff (95% CI)	---	-11.9 (-30.3, 6.4)	16.0 (-7.5, 39.5)
Supporting Analysis: Using ANCOVA method				
LOCF in FAS	LS Mean (95% CI)	-167.0 (-184.5, -149.4)	-169.8 (-184.0, -155.6)	-147.9 (-165.7, -130.1)
	Diff (95% CI)	---	-2.8 (-21.2, 15.6)	19.1 (-2.1, 40.4)

OC: Observed cases; LOCF: Last observation carried forward; MMRM: Mixed model repeated measurements; FAS: Full analysis set.
Source: Reviewer's analysis based on ADMOOCT.xpt dataset located at [\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\vgftehd-dme-1934\analysis\adam\datasets](#) for PHOTON.

Figure 14 shows the least squares mean changes in CST over time, based on the observed cases analyzed with MMRM. As shown, the adjusted means decreased from baseline in CST through visits, in cyclic trajectories among treatment groups.

Figure 14: Least Square Mean Changes from Baseline in CST (in μ m) by visit, MMRM, FAS (PHOTON)



Source: Based on Reviewer's analysis.

Reviewer's Comment:

Figure 14 above sheds some light on the impact of frequent dosing versus less frequent dosing on the CST outcome over time. As shown, up to Week 12 where all groups received fixed monthly dosing, a comparable and smooth continued reduction in CST was observed in all groups. However, a wave-like pattern and a shift in these patterns observed beyond Week 12 in each treatment group was likely attributed to the less frequent dosing. For example, beyond Week 12, the reduction in CST worsened up until the next injection visit due to extended dosing (e.g., at Week 20 for HDq12, and at Week 24 for 2q8 and HDq16) and then a better reduction was observed in the next immediate visit after the injection (e.g., at Week 24 in HDq12, and at Week 28 in 2q8 and HDq16). The same pattern continued through Week 48.

3.2.2.5. Efficacy Conclusion

Based on the efficacy evidence from the PHOTON study, the reviewer concludes that subjects treated with aflibercept 8 mg administered every 12 weeks (HDq12) and every 16 weeks (HDq16) after three initial monthly doses had a noninferior mean change in BCVA from baseline at Week 48 compared to subjects treated with aflibercept 2 mg administered every 8 weeks after five initial monthly dosing (2q8). It should be noted that although both HDQ12 and HDq16 were noninferior to 2q8 in the primary endpoint, subjects treated with 2q8 displayed numerically better BCVA gain at each visit through Week 48 compared to subjects treated with HDq12 and HDq16 (*Figure 9*).

Additionally, in the PHOTON study, HDq12 was noninferior to 2q8 in the key secondary endpoint of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline at Week 48. However, noninferiority could not be shown for the HDq16 group because the lower bound of the 95% CI for the difference in the adjusted proportions between HDq16 and 2q8 was less than the noninferiority margin of -10%. Thus, for the DR indication, only HDq12 provided comparable efficacy to 2q8.

3.3. EVALUATION OF SAFETY

In PULSAR study for nAMD indication and in the PHOTON study for the DME/DR indication, safety was evaluated based on the safety evaluable population including all randomized subjects who received at least a single dose of double-blind treatment. The safety parameters included extent of exposure to study drug, adverse events (AEs), clinical laboratory evaluation, and additional safety variables which included intraocular pressure, vital signs, electrocardiogram, and ophthalmic examinations.

In this review, a high-level safety summary is provided for each study by indication. For a comprehensive safety evaluation, the reviewer defers to the FDA clinical review document.

3.3.1. **PULSAR Study**

Study Exposure

Table 31 shows summary of study treatment exposure through Week 48 in the PULSAR study. In the study, most of the randomized subjects received at least one dose of study treatment. Two randomized subjects (1 each from 2q8 and HDq12) who did not receive any treatment were excluded from the safety-evaluable population. As shown, the median duration of exposure through Week 48 in the PULSAR study was comparable between the treatment groups (48 weeks). The median number of active injections and the total number of injections most subjects received through Week 48 were 7, 6, and 5 in the 2q8, HDq12, and HDq16 groups, respectively.

Table 31: Summary of Study Treatment Exposure in the Study Eye through Week 48 (Safety Analysis Set) (PULSAR)

	2q8 (N = 336)	HDq12 (N = 335)	HDq16 (N = 338)	All HD (N = 673)
Treatment Duration				
N	336	335	337	672
Mean (SD)	46.3 (6.62)	46.5 (6.56)	46.2 (6.97)	46.4 (6.77)
Median	48.0	48.0	48.0	48.0
Range	4 – 50.9	4 – 52	4 – 53.3	4 – 53.3
Number of Active Injections				
N	336	335	337	672
Mean (SD)	6.7 (0.8)	5.9 (0.8)	5.1 (0.8)	5.5 (0.9)
Median	7.0	6.0	5.0	6.0
Range	1 - 8	1 - 7	1 - 7	1 - 7
Number of Active Injections				
1	1 (0.3%)	2 (0.6%)	2 (0.6%)	4 (0.6%)
2	1 (0.3%)	2 (0.6%)	1 (0.3%)	3 (0.4%)
3	4 (1.2%)	3 (0.9%)	9 (2.7%)	12 (1.8%)
4	6 (1.8%)	7 (2.1%)	22 (6.5%)	29 (4.3%)
5	6 (1.8%)	22 (6.6%)	263 (77.8%)	285 (42.3%)
6	29 (8.6%)	260 (77.6%)	11 (3.3%)	271 (40.3%)
7	288 (85.7%)	39 (11.6%)	29 (8.6%)	68 (10.1%)
8	1 (0.3%)	0	0	0

Source: Table 10-1 of Applicant's Study Report

In the PULSAR study, a total of 628 subjects (308 in 2q8, 316 in HDq12, and 312 in HDq16) completed the study through Week 48. Of the subjects in the HDq12 and HDq16 arms that completed the study through Week 48, (i) a total of 523 (83%) of subjects assigned in the HDq12 or HDq16 arm maintained q12 or longer dosing interval, (ii) a total of 251 (79%) subjects assigned in the HDq12 arm maintained the q12 dosing interval, and (iii) 239 (77%) subjects assigned in the HDq16 arm maintained the q16 dosing interval. Sixty-five (65) subjects in HDq12 arm shortened to q8 dosing interval and 73 subjects in the HDq16 arm shortened to either q8 interval (n = 40) or q12 interval (n = 33).

Adverse Events:

In the PULSAR study, a total of 1009 subjects were exposed to the study drug through Week 48 (336 in 2q8 and 673 in the pooled high dose of aflibercept).

A high-level safety summary of the study is presented in [Table 32](#). As shown, the incidence of overall Treatment Emergent Adverse Events (TEAEs), ocular TEAEs, serious TEAEs, ocular serious TEAEs, and TEAEs of intraocular inflammation (IOI) were comparable between the treatment groups. The non-ocular TEAEs and serious TEAEs were also comparable between the treatment groups.

A comparable number of subjects across the treatment groups discontinued the study drug due to any TEAEs.

Table 32: Overview of Safety Through Week 48 from PULSAR Study (Safety Analysis Set)

	2q8 (N = 336)	HDq12 (N = 335)	HDq16 (N = 338)	All HD (N = 673)
Any AE	240 (71.4%)	242 (72.2%)	254 (75.1%)	496 (73.7%)
Any TEAE	235 (69.9%)	239 (71.3%)	249 (73.7%)	488 (72.5%)
Ocular: Study Eye	130 (38.7%)	129 (38.5%)	127 (37.6%)	256 (38.0%)
Non-ocular	178 (53.0%)	175 (52.2%)	182 (53.8%)	357 (53.0%)
Any Serious TEAE	49 (14.6%)	41 (12.2%)	37 (10.9%)	78 (11.6%)
Ocular: Study Eye	2 (0.6%)	6 (1.8%)	5 (1.5%)	11 (1.6%)
Non-ocular	46 (13.7%)	34 (10.1%)	32 (9.5%)	66 (9.8%)
Any TEAE related to intravitreal injection procedure	37 (11.0%)	37 (11.3%)	33 (9.8%)	71 (10.5%)
Ocular: Study Eye	35 (10.4%)	32 (9.6%)	32 (9.5%)	64 (9.5%)
Non-ocular	0	4 (1.2%)	1 (0.3%)	5 (0.7%)
Any Serious TEAE related to intravitreal injection procedure	0	3 (0.9%)	2 (0.6%)	5 (0.7%)
Ocular: Study Eye	0	2 (0.6%)	2 (0.6%)	4 (0.6%)
Non-ocular	0	1 (0.3%)	0	1 (0.1%)
Any TEAE of IOI in the study eye	2 (0.6%)	4 (1.2%)	1 (0.3%)	5 (0.7%)
Any TEAE leading to study drug discontinuation	5 (1.5%)	5 (1.5%)	5 (1.5%)	10 (1.5%)
Any TEAE with outcome death	5 (1.5%)	3 (0.9%)	1 (0.3%)	4 (0.6%)
Any Adjudicated Treatment-Emergent APTC Events	5 (1.5%)	1 (0.3%)	1 (0.3%)	2 (0.3%)
Cardiovascular (CV) death	1 (0.3%)	0	0	0
Non-fatal myocardial infarction	2 (0.6%)	0	1 (0.3%)	1 (0.1%)
Non-fatal stroke	2 (0.6%)	1 (0.3%)	0	1 (0.1%)

Source: Table 10-3 and Table 10-16 of Applicant's Clinical Study Report

A total of nine deaths were reported in the PULSAR study through Week 48 due to any TEAEs: five deaths in 2q8 and four deaths in the pooled high dose of aflibercept (3 in HDq12 and 1 in HDq16).

In the PULSAR study, adjudicated arterial thromboembolic events based on Anti-Platelet Trialists' Collaboration (APTC) Endpoint were assessed. The APTC events included a non-fatal strokes or non-fatal myocardial infarctions or vascular deaths (including deaths of unknown cause). The summary of these events is presented in [Table 32](#). As shown, slightly more 2q8 treated subjects had adjudicated APTC in the study.

In summary, the overall safety profile across the treatment groups in the PULSAR study appear comparable.

3.3.2. **PHOTON Study**

[Table 33](#) shows the summary of study treatment exposure through Week 48 in the PHOTON study. In the study, most of the randomized subjects received at least one dose of study treatment. Two randomized subjects (1 each from HDq12 and HDq16) who did not receive any treatment were excluded from the safety-evaluable population. As shown, the median duration of exposure through Week 48 in the PHOTON study was comparable between the treatment groups (48 weeks). The median number of active injections and the total number of injections most subjects received through Week 48 were 8, 6, and 5 in the 2q8, HDq12, and HDq16 groups, respectively.

Table 33: Summary of Study Treatment Exposure in the Study Eye through Week 48 (Safety Analysis Set) (PHOTON)

	2q8 (N = 167)	HDq12 (N = 328)	HDq16 (N = 163)	All HD (N = 491)
Treatment Duration				
N	167	328	163	491
Mean (SD)	46.7 (6.89)	45.6 (9.0)	47.1 (6.0)	46.1 (8.2)
Median	48.0	48.0	48.0	48.0
Range	4 – 52.1	4 – 64.4	4 – 54.9	4 – 64.4
Number of Active Injections				
N	167	328	163	491
Mean (SD)	7.7 (0.98)	5.7 (0.96)	4.9 (0.61)	5.5 (0.93)
Median	8.0	6.0	5.0	6.0
Range	1 - 8	1 - 7	1 - 7	1 - 7
Number of Active Injections				
1	2 (1.2%)	2 (0.6%)	1 (0.6%)	3 (0.6%)
2	0	7 (2.1%)	1 (0.6%)	8 (1.6%)
3	0	11 (3.4%)	3 (1.8%)	14 (2.9%)
4	2 (1.2%)	11 (3.4%)	6 (3.7%)	17 (3.5%)
5	2 (1.2%)	12 (3.7%)	146 (89.6%)	158 (32.2%)
6	3 (1.8%)	273 (83.2%)	2 (1.2%)	275 (56%)
7	10 (6.0%)	12 (3.7%)	4 (2.5%)	16 (3.3%)
8	148 (88.6%)	0	0	0

Source: Table 10-1 of Applicant's Study Report

In the PHOTON study, a total of 613 subjects (157 in 2q8, 300 in HDq12, and 156 in HDq16) completed the study through Week 48. Of the subjects in the HDq12 and HDq16 arms that completed the study through Week 48, (i) a total of 413 (93%) of subjects assigned in the HDq12 or HDq16 arm maintained q12 or longer dosing interval, (ii) a total of 273 (91%) subjects assigned in the HDq12 arm maintained the q12 dosing interval, and (iii) 139 (89%) subjects assigned in the HDq16 arm maintained the q16 dosing interval. Twenty-seven (27) subjects in HDq12 arm shortened to q8 dosing interval and 17 subjects in the HDq16 arm shortened to either q8 interval (n = 6) or q12 interval (n = 13).

Adverse Events:

In the PHOTON study, a total of 658 subjects were exposed to the study drug through Week 48 (167 in 2q8 and 491 in the pooled high dose of aflibercept).

A high-level safety summary of the study is presented in [Table 34](#). As shown, the incidence of overall TEAEs, ocular TEAEs, serious TEAEs, ocular serious TEAEs, and TEAEs of IOI were comparable between the treatment groups. The non-ocular TEAEs and serious TEAEs were also comparable between the treatment groups. A comparable number of subjects across the treatment groups discontinued the study drug due to any TEAEs.

Table 34: Overview of Safety Through Week 48 from PULSAR Study (Safety Analysis Set)

	2q8 (N = 167)	HDq12 (N = 328)	HDq16 (N = 163)	All HD (N = 491)
Any AE				
Any TEAE	106 (63.5%)	227 (69.2%)	118 (72.4%)	345 (70.3%)
Ocular: Study Eye	46 (27.5%)	104 (31.7%)	48 (29.4%)	152 (31.0%)
Non-ocular TEAE	79 (47.3%)	174 (53.0%)	95 (58.3%)	269 (54.8%)
Any Serious TEAE	31 (18.6%)	55 (16.8%)	24 (14.7%)	79 (16.1%)
Ocular: Study Eye	1 (0.6%)	2 (0.6%)	1 (0.6%)	3 (0.6%)
Non-ocular	26 (15.6%)	52 (15.9%)	22 (13.5%)	74 (15.1%)
Any TEAE related to intravitreal injection procedure	2 (1.2%)	6 (1.8%)	0	6 (1.2%)
Ocular: Study Eye	0	2 (0.6%)	0	2 (0.4%)
Non-ocular TEAE	2 (1.2%)	4 (1.2%)	0	4 (0.8%)
Any Serious TEAE related to intravitreal injection procedure	0	1 (0.3%)	0	1 (0.2%)
Ocular: Study Eye	0	1 (0.3%)	0	1 (0.2%)
Non-ocular TEAE	0	0	0	0
Any TEAE of IOI in the study eye	1 (0.6%)	4 (1.2%)	0	4 (0.8%)
Any TEAE leading to study drug discontinuation	2 (1.2%)	8 (2.4%)	2 (1.2%)	10 (2.0%)
Any TEAE with outcome death	4 (2.4%)	8 (2.4%)	2 (1.2%)	10 (2.0%)
Any Adjudicated Treatment-Emergent APTC Events	6 (3.6%)	8 (2.4%)	7 (4.3%)	15 (3.1%)
Cardiovascular (CV) death	3 (1.8%)	2 (0.6%)	1 (0.6%)	3 (0.6%)
Non-fatal myocardial infarction	3 (1.8%)	3 (0.9%)	1 (0.6%)	4 (0.8%)
Non-fatal stroke	0	3 (0.9%)	5 (3.1%)	8 (1.6%)

Source: Table 27 and Table 34 of Applicant's Clinical Study Report

A total of 14 deaths were reported in the PHOTON study through Week 48 due to any TEAEs: four deaths in 2q8 and 10 deaths in the pooled high dose of aflibercept (eight in HDq12 and two in HDq16).

In the PHOTON study, adjudicated arterial thromboembolic events based on APTC Endpoint were assessed. The APTC events included a non-fatal strokes or non-fatal myocardial infarctions or vascular deaths (including deaths of unknown cause). The summary of these events is presented in [Table 34](#). As shown, a comparable number of subjects in each treatment group had adjudicated APTC in the study.

In summary, the overall safety profile across the treatment groups in the PHOTON study appear comparable.

4. Findings in Special/Subgroup Populations

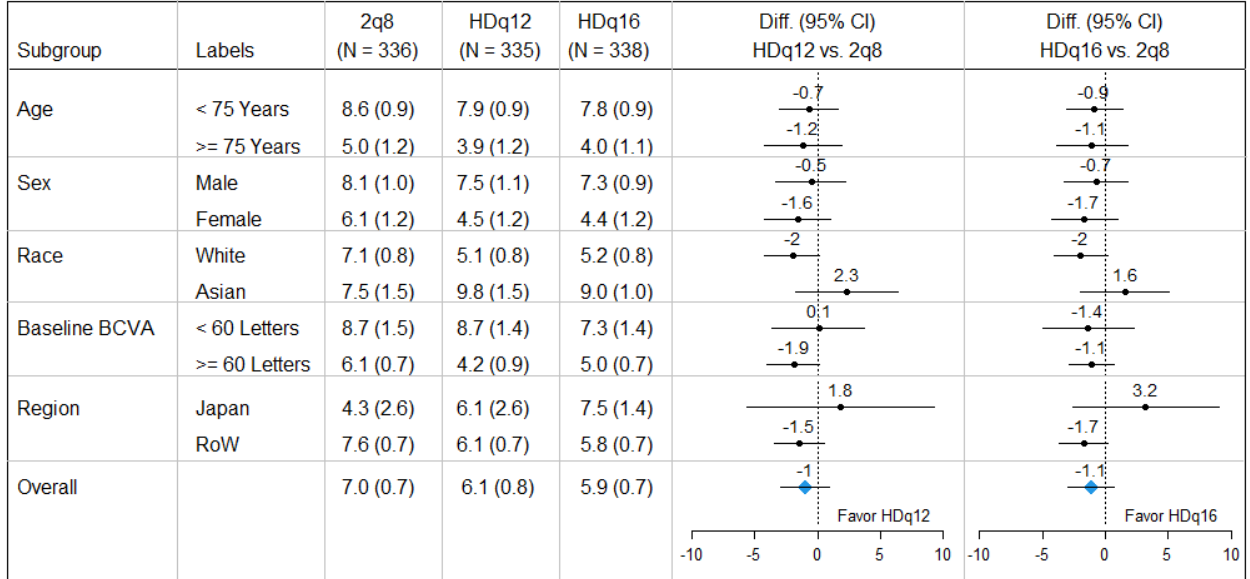
In this section, the primary efficacy variable of the change in BCVA from baseline at Week 48 in the PULSAR study for the nAMD indication and in the PHOTON study for the DME indication, and the key secondary efficacy variable of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline at Week 48 in the PHOTON study for DR indication were summarized by subgroup variables.

4.1. PULSAR STUDY FOR NAMD INDICATION

In the PULSAR study, the primary efficacy variable of the change in BCVA from baseline at Week 48 was summarized by the subgroups of age, sex, and race, and by the stratification factors of baseline BCVA category and region. It should be noted that subgroup summary was not reported for some categories of race due to small sample sizes. See Table 6 for the sample sizes used within each subgroup.

Figure 15 below displays the adjusted mean change in BCVA from baseline at Week 48 by the subgroup variables in the PULSAR study. As shown, within the levels of each of the subgroup variables, the average number of letters gained from baseline at Week 48 was comparable across the treatment groups. Additionally, the adjusted differences in the mean change in BCVA at Week 48 between the treatment groups across the subgroup levels were consistent with the overall population.

Figure 15: Adjusted Mean Change in BCVA from Baseline at Week 48 by Subgroup (PULSAR) (Full Analysis Set)



Source: Based on Reviewer’s analysis.

For some subgroups since there were a small number of subjects, the results for these subgroups may not be indicative of the overall treatment effects.

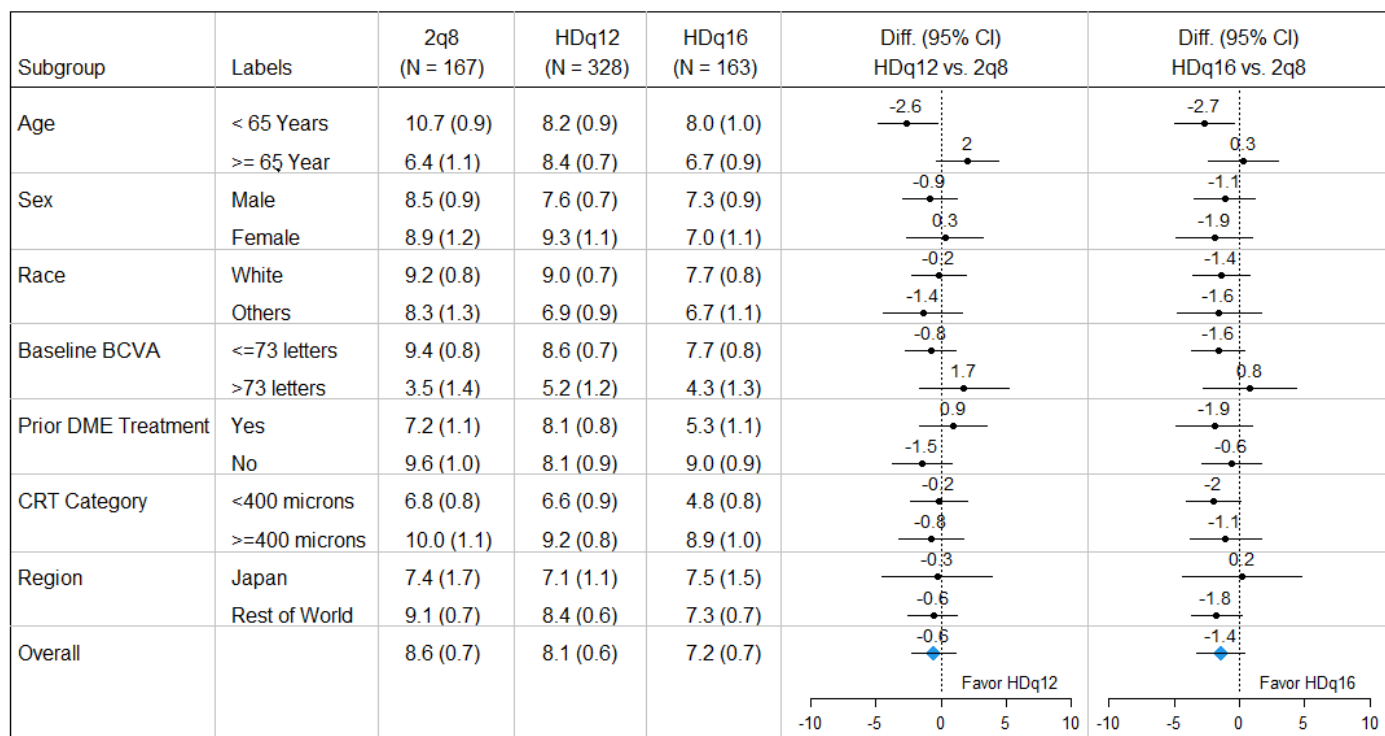
4.2. PHOTON STUDY FOR DME/DR INDICATIONS

In the PHOTON study, the primary efficacy variable of the change in BCVA from baseline at Week 48 was summarized by the subgroups of age, sex, and race, baseline BCVA category, and by the stratification factors of prior DME treatment use, CRT category, and region. It should be noted that subgroup summary was not reported for some categories of race due to small sample sizes. See Table 18 for the sample sizes used within each subgroup.

Figure 16 below displays the summary of the mean change in BCVA from baseline at Week 48 by the subgroup variables in the PHOTON study. As shown, the average number of letters gained from baseline at Week 48 within the levels of each of the subgroup variables was comparable across the treatment groups. Additionally, the adjusted differences in the mean change in BCVA at Week 48 between the treatment groups across the subgroup levels were consistent with the overall population.

For some subgroups since there were a small number of subjects, the results for these subgroups may not be indicative of the overall treatment effects.

Figure 16: Adjusted Mean Change in BCVA from Baseline at Week 48 by Subgroup (PHOTON)
(Full Analysis Set)



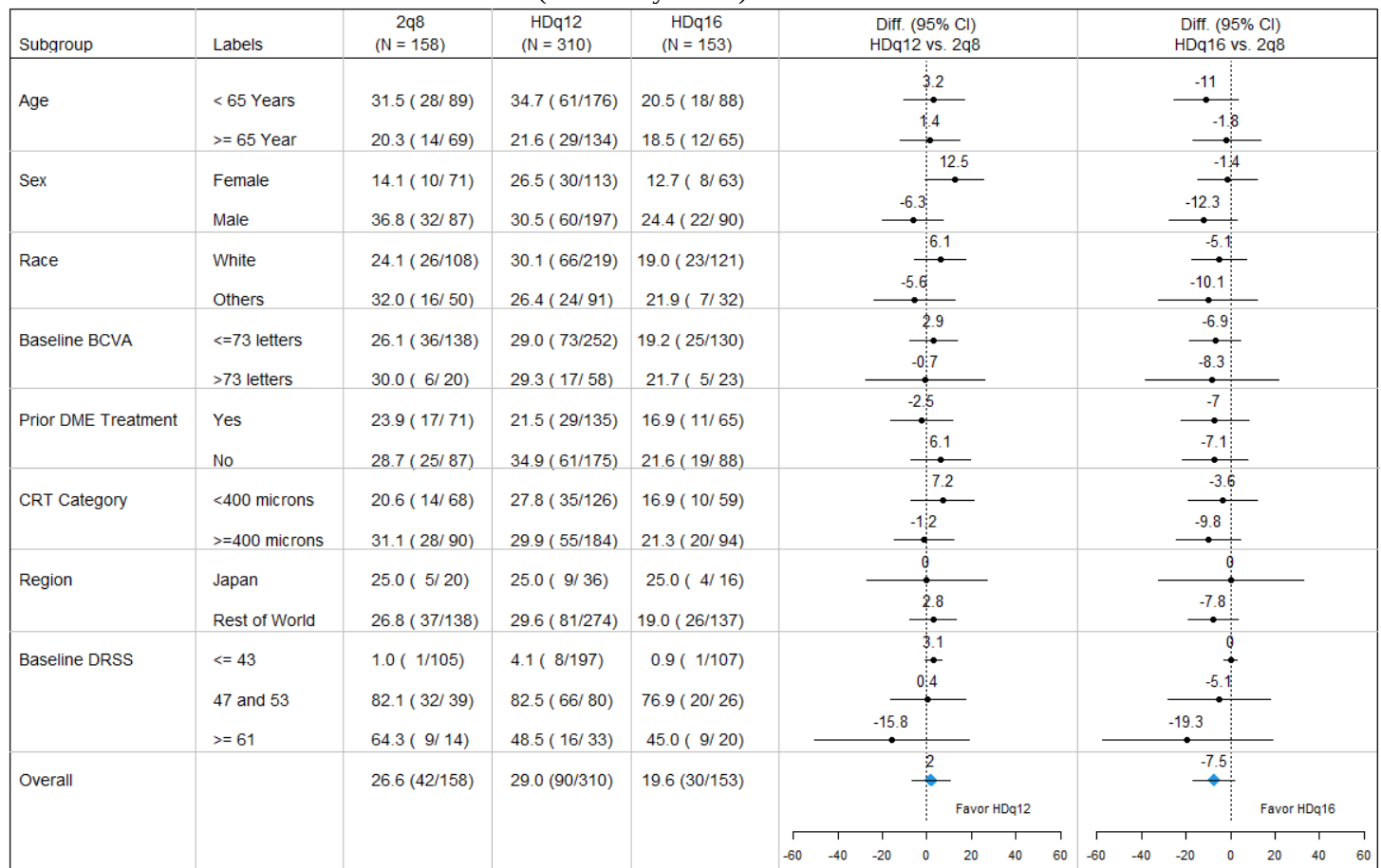
Source: Based on Reviewer's analysis.

In the PHOTON study, the key secondary efficacy variable of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline at Week 48 was also summarized by the subgroup variables of age, sex, race, baseline BCVA category, and by the stratification factors of prior DME treatment use, CRT category, and region.

Figure 17 below displays the summary of the proportion of subjects who achieved ≥ 2 -step improvement in DRSS from baseline at Week 48 by the levels of the subgroup variables including the treatment differences (denoted by dots) and 95% confidence interval estimates (denoted by horizontal bars).

As shown, the results of the key secondary efficacy variable within the levels of each of the subgroup variables, except in subjects with mild PDR or worse at baseline (DRSS ≥ 61), were consistent with those in the overall population. In the PHOTON study, subject with baseline DRSS ≥ 61 appeared to have benefitted with more frequent dosing with 2q8 than with less frequent dosing of HDq12 or HDq16. See Section 3.2.2.2 for a detailed efficacy assessment of the key secondary endpoint by baseline DRSS category.

Figure 17: Proportion of Subjects Who Achieved ≥ 2 -Step Improvement in DRSS from Baseline at Week (PHOTON)48 by Subgroup (Full Analysis Set)



Source: Based on Reviewer's analysis.

5. Summary and Conclusions

5.1. STATISTICAL ISSUES

There are no major statistical issues in this application. However, the following issue was identified and addressed during the review.

PHOTON study underwent four protocol amendments. Up until protocol amendment 3, the Applicant intended to use a graphical based alpha propagation strategy to control the study-wise Type I error rate at a two-sided significance level of 5% (Figure 8). According to the graph-based testing strategy, noninferiority in the key secondary DR-related efficacy endpoints would only have been tested if noninferiority of HDq12 to 2q8 and HDq16 to 2q8 in the primary efficacy endpoint and superiority in dryness endpoint at Week 12 were established. In protocol amendment 4, the Applicant removed the first key secondary endpoint of dryness in the testing hierarchy one month before the last patient-last visit date and four months before the database lock.

The Applicant indicated that the decision to remove the endpoint was based on “*recent data from another, completed study*”. However, such data and the completed study were not shared with the Agency at the time of the BLA submission. Thus, considering that the removed endpoint showed a negative result (Table 29), the Agency questioned the timing of such major protocol change and the reliability of testing the DR-related endpoints in the hierarchy.

The Applicant in this case provided a reasonable justification for their decision to remove the endpoint in the hierarchy (see Section 3.2.2.2 [RE: Type I Error Control (Plan for Multiplicity Adjustment)]) but the Reviewer considers that a major protocol change such as this near the end of the study should be discouraged in general.

5.2. COLLECTIVE EVIDENCE

Support for the efficacy and safety of aflibercept 8 mg for the treatment of nAMD and DME/DR was based on two pivotal studies designed to assess whether aflibercept 8 mg administered every 12 or every 16 weeks reduce the treatment burden while maintaining comparable efficacy benefit to the approved aflibercept 2 mg every 8 weeks.

For the nAMD indication, the primary efficacy evidence to support noninferiority of aflibercept 8 mg administered every 12 or 16 weeks after three initial monthly doses to aflibercept 2 mg administered every 8 weeks after three initial monthly doses was based on one pivotal phase 3 trial (PULSAR study). Subjects treated in the high dose groups (HDq12 and HDq16) had a noninferior mean change in BCVA from baseline at Week 48 compared to subjects treated with aflibercept 2 mg in the 2q8 group because the 95% lower confidence interval limit for the treatment differences in the adjusted means were greater than the noninferiority margin of -4.0 letters (Table 1). Although noninferiority in the primary endpoint was established in the study, because of the frequent dosing, subjects that received aflibercept 2q8 displayed numerical advantage in the BCVA gain from baseline through Week 48 compared to subjects that received the high doses of aflibercept (Figure 3).

For the DME indication, the primary efficacy evidence to support noninferiority of aflibercept 8 mg administered every 12 or 16 weeks after three initial monthly doses to aflibercept 2 mg administered every 8 weeks after five initial monthly doses was based on one pivotal phase 2/3 trial (PHOTON study). In this study, subjects treated with high dose of aflibercept administered in every 12-week and every 16-week dosing regimen had a noninferior mean change in BCVA from baseline at Week 48 compared to subjects treated with aflibercept 2 mg administered in every 8-week dosing regimen because the 95% lower confidence interval limit for the treatment differences in the adjusted means was greater than the noninferiority margin of -4.0 letters ([Table 2](#)). Although noninferiority in the primary endpoint was established in the study, because of the frequent dosing, subjects that received aflibercept 2q8 displayed numerical advantage in the BCVA gain from baseline through Week 48 compared to subjects that received the high doses of aflibercept ([Figure 9](#)).

Support for the DR indication in PHOTON study was based on the key secondary efficacy endpoint of the proportion of subjects who achieved a ≥ 2 -step improvement in DRSS from baseline at Week 48. In this study, HDq12 was noninferior to 2q8 in the key secondary endpoint. However, noninferiority could not be shown for the HDq16 group because the lower bound of the 95% CI for the difference in the adjusted proportions between HDq16 and 2q8 was less than the noninferiority margin of -10%. Thus, for the DR indication, only HDq12 provided comparable efficacy to 2q8.

5.3. CONCLUSION AND RECOMMENDATION

Based on the collective efficacy evidence from the adequate and well controlled PULSAR study, the reviewer concludes that the application for the nAMD indication provided substantial evidence for comparable efficacy benefit of aflibercept 8 mg administered every 12 or 16 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after three initial monthly doses.

Similarly, based on the collective efficacy evidence from the adequate and well controlled PHOTON study, the reviewer concludes that the application for the DME indication provided substantial evidence for comparable efficacy benefit of aflibercept 8 mg administered every 12 or 16 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after five initial monthly doses.

For the DR indication, the reviewer concludes that the application based on the PHOTON study provided substantial evidence for comparable efficacy of aflibercept 8 mg administered every 12 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after five initial doses. On the other hand, for this indication, the study did not establish substantial evidence for comparable efficacy of aflibercept 8 mg administered every 16 weeks after three initial monthly doses compared to aflibercept 2 mg administered every 8 weeks after 5 initial monthly doses.

5.4. LABELING RECOMMENDATION

The Applicant's proposed texts, tables, and figures presented in Section 14.1 (for the nAMD indication), in Section 14.2 (for the DME indication), and in Section 14.3 (for DR indication) of the draft label seem reasonable.

Regarding Table 3 and Figure 8 presented in Section 14.1 for nAMD indication, the reviewer has the following recommendations (highlighted in red font):

Table 3: Efficacy Outcomes (Full Analysis Set) in PULSAR Study

Efficacy Outcomes	HDq12	HDq16	2q8
Full Analysis Set ^a	N=335	N=338	N=336
Mean change in BCVA as measured by ETDRS letter score from baseline (SD) at week 48 ^b	6.7 (12.6)	6.2 (11.7)	7.6 (12.2)
LS mean (SE) change from baseline ^c	6.1 (0.8)	5.9 (0.7)	7.0 (0.7)
Difference in LS mean (95% CI) ^c	-1.0 (-2.9, 0.9)	-1.1 (-3.0, 0.7)	

BCVA = Best Corrected Visual Acuity; ETDRS – Early Treatment Diabetic Retinopathy Study; SD = Standard Deviation; LS = Least Square; SE = Standard Error; CI = Confidence Interval; MMRM = Mixed Model for Repeated Measurements

^a Full Analysis Set (FAS) includes all randomized patients who received at least 1 dose of study medication

^b Observed values at week 48: n=299 for HDq12; n=289 for HDq16; n=285 for 2q8

^c Estimate based on the MMRM model, was computed for the differences of HDq12 minus 2q8 and HDq16 minus 2q8, respectively, with two-sided 95% CIs

- i) To minimize redundancy and confusion, only the adjusted means in the change in BCVA be presented in the table.

iv) Figure 8 should be re-produced based on the adjusted means.

Regarding Table 4 and Figure 9 presented in Section 14.2 for the DME indication, the reviewer has the following recommendations (highlighted in red font):

Table 4: Efficacy Outcomes (Full Analysis Set) in PHOTON Study

Efficacy Outcomes	HDq12	HDq16	2q8
Full Analysis Set ^a	N=328	N=163	N=167
Mean change in BCVA as measured by ETDRS letter score from baseline (SD) at week 48 ^b	8.8 (9.0)	7.9 (8.4)	9.2 (9.0)
LS mean (SE) change from baseline ^c	8.1 (0.6)	7.2 (0.7)	8.7 (0.7)
Difference in LS mean (95% CI) ^c	-0.6 (-2.3, 1.1)	-1.4 (-3.3, 0.4)	

(b) (4)

BCVA = Best Corrected Visual Acuity; ETDRS = Early Treatment Diabetic Retinopathy Study; SD = Standard Deviation; LS = Least Square; SE = Standard Error; CI = Confidence Interval.

^a FAS includes all randomized patients who received at least 1 dose of study medication

^b Observed values at week 48: n=277 for HDq12; n=149 for HDq16; n=150 for 2q8

^c Estimate based on the MMRM model, was computed for the differences of HDq12 minus 2q8 and HDq16 minus 2q8, respectively with two-sided 95% CIs

i) To minimize redundancy and confusion, only the adjusted means in the change in BCVA be presented in the table.

(b) (4)

iii) Figure 9 should be re-produced based on the adjusted means.

Regarding Table 5, Figure 10, and Table 6 presented in Section 14.3 for the DR indication, the reviewer has the following recommendations (highlighted in red font):

i) Remove HDq16 column from Table 6 because noninferiority of HDq16 to 2q8 was not established in the study.

Table 5: Proportion of Patients Who Achieved a ≥ 2 -Step Improvement from Baseline in the ETDRS-DRSS Score at Week 48 (Full Analysis Set) in PHOTON

Efficacy Outcomes	HDq12	~~HDq16~~	2q8
Full Analysis Set ^a	N=328	~~N=163~~	N=167
Proportion of patients with a ≥ 2 -step improvement on ETDRS-DRSS from Baseline (%) ^b	29%	~~20%~~	27%
Difference ^c (%) (95% CI)	2% (-6.6, 10.6)	~~-8%~~ ~~(-16.9, -1.8)~~	

Missing or non-gradable post-baseline ETDRS-DRSS values were imputed using the last gradable ETDRS-DRSS values. Patients were considered as non-responders if all post-baseline measurements were missing or non-gradable. Missing or ungradable baseline was not included in the denominator.

- ^a FAS includes all randomized patients who received at least 1 dose of study medication
- ^b Last observation carried forward
- ^c Difference with confidence interval (CI) was calculated using Mantel-Haenszel weighting scheme

ii) [REDACTED] (b) (4)

iii) [REDACTED] (b) (4)

For the DR indication, in addition to the recommended changes in Section 14.3, the DOSAGE and ADMINISTRATION section should also be revised as follows (see red font) [REDACTED] (b) (4)

- **Diabetic Retinopathy (DR)**

- The recommended dose for EYLEA HD is 8 mg (0.07 mL [REDACTED] (b) (4)) administered by intravitreal injection every 4 weeks (approximately every 28 days +/- 7 days, [REDACTED] (b) (4)) for the first three doses, followed by 8 mg (0.07 mL) via intravitreal injection once every 8 – [REDACTED] (b) (4) 12 weeks [REDACTED] (b) (4) +/- 7 days) (2.4)

APPENDICES

Appendix 1: PULSAR Study

Table 35: Adjusted Mean Change in BCVA from Baseline Through Week 48 by Treatment (PULSAR)

	Difference in LS Means (95% CI)				
	2q8	HDq12	HDq16	HDq12 vs 2q8	HDq16 vs. 2q8
Week 4	4.2 (0.5)	2.4 (0.5)	2.9 (0.5)	-1.8 (-2.9, -0.6)	-1.2 (-2.3, -0.1)
Week 8	5.9 (0.6)	4.5 (0.6)	4.4 (0.5)	-1.4 (-2.7, -0.0)	-1.4 (-2.7, -0.2)
Week 12	6.2 (0.6)	5.0 (0.6)	5.5 (0.6)	-1.2 (-2.7, 0.3)	-0.7 (-2.2, 0.7)
Week 16	6.2 (0.6)	5.5 (0.6)	6.1 (0.6)	-0.7 (-2.2, 0.7)	-0.1 (-1.5, 1.3)
Week 20	7.1 (0.7)	5.6 (0.6)	5.9 (0.6)	-1.5 (-3.1, 0.0)	-1.2 (-2.7, 0.3)
Week 24	7.1 (0.7)	5.4 (0.7)	5.4 (0.7)	-1.7 (-3.4, -0.0)	-1.7 (-3.3, -0.1)
Week 28	7.7 (0.7)	6.5 (0.7)	5.8 (0.7)	-1.2 (-2.8, 0.5)	-1.9 (-3.5, -0.2)
Week 32	7.2 (0.7)	6.2 (0.7)	6.8 (0.7)	-0.9 (-2.6, 0.8)	-0.3 (-2.0, 1.3)
Week 36	7.7 (0.7)	6.1 (0.7)	5.9 (0.7)	-1.6 (-3.4, 0.2)	-1.8 (-3.7, -0.0)
Week 40	7.2 (0.7)	6.4 (0.7)	5.4 (0.7)	-0.8 (-2.6, 1.0)	-1.9 (-3.7, -0.1)
Week 44	7.4 (0.7)	6.7 (0.7)	5.8 (0.7)	-0.7 (-2.4, 1.1)	-1.6 (-3.4, 0.1)
Week 48	7.0 (0.7)	6.1 (0.8)	5.9 (0.7)	-1.0 (-2.9, 0.9)	-1.1 (-3.0, 0.7)

Source: Based on Reviewer's analysis

Table 36: Change from baseline in BCVA at Week 48 - Multiple Imputation with Tipping Point Analysis (PULSAR)
(FAS)

Shift value	LS Means (SE)			Difference in LS Means (95% CI)	
	2q8	HDq12	HDq16	HDq12 vs. 2q8	HDq16 vs. 2q8
0	7.2 (0.87)	6.2 (0.87)	6.1 (0.85)	-0.9 (-2.8, 0.9)	-1.1 (-3.0, 0.8)
-1	7.2 (0.87)	6.1 (0.87)	5.9 (0.85)	-1.0 (-2.9, 0.8)	-1.2 (-3.1, 0.6)
-2	7.2 (0.87)	6.1 (0.87)	5.8 (0.85)	-1.1 (-3.0, 0.7)	-1.4 (-3.3, 0.5)
-3	7.2 (0.87)	6.0 (0.87)	5.7 (0.86)	-1.2 (-3.1, 0.6)	-1.5 (-3.4, 0.4)
-4	7.2 (0.88)	5.9 (0.87)	5.6 (0.86)	-1.4 (-3.2, 0.5)	-1.7 (-3.5, 0.2)
-5	7.3 (0.88)	5.8 (0.88)	5.4 (0.86)	-1.5 (-3.4, 0.4)	-1.8 (-3.7, 0.1)
-6	7.3 (0.88)	5.7 (0.88)	5.3 (0.86)	-1.6 (-3.5, 0.3)	-2.0 (-3.8, -0.1)
-7	7.3 (0.89)	5.6 (0.88)	5.2 (0.87)	-1.7 (-3.6, 0.2)	-2.1 (-4.0, -0.2)
-8	7.3 (0.89)	5.5 (0.89)	5.1 (0.87)	-1.8 (-3.7, 0.1)	-2.2 (-4.2, -0.3)
-9	7.3 (0.89)	5.4 (0.89)	4.9 (0.88)	-1.9 (-3.8, 0.0)	-2.4 (-4.3, -0.5)
-10	7.3 (0.90)	5.4 (0.90)	4.8 (0.88)	-2.0 (-3.9, -0.0)	-2.5 (-4.5, -0.6)
-11	7.4 (0.90)	5.3 (0.90)	4.7 (0.89)	-2.1 (-4.0, -0.1)	-2.7 (-4.6, -0.7)
-12	7.4 (0.91)	5.2 (0.91)	4.6 (0.89)	-2.2 (-4.2, -0.2)	-2.8 (-4.8, -0.9)

LS Means: Least square Means; SE: Standard error; CI: Confidence Interval

Shift value = 0 represents results based on the standard multiple imputation (MI) approach without a shift.

Shift value -1 to -12 represent Tipping Point analysis results.

Using a noninferiority (NI) margin of -4 letters, NI was established for shift value = 0. This conclusion of noninferiority was reversed at shift value -11 for HDq12 vs. 2q8 and at shift value -7 for HDq16 vs. 2q8.

Table 37: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 16. (Sensitivity and Supportive Analyses) (PULSAR)

	2q8	HDq12	HDq16	All HD
OC, FAS				
Participants who had no IRF and no SRF, Num/Den (%)	166/326 (50.9%)	197/325 (60.6%)	211/325(64.9%)	408/650 (62.8%)
Difference % (95% CI)				11.9 (5.3, 18.5)
LOCF, PPS				
Participants who had no IRF and no SRF, Num/Den (%)	161/320 (50.3%)	197/324 (60.8%)	208/324 (64.2%)	405/648 (62.5%)
Difference % (95% CI)				12.2 (5.6, 18.8)

OC: Observed cases; LOCF: Last observation carried forward; FAS: Full analysis set; PPS: Per protocol set; IRF: intraretinal fluid; SRF: subretinal fluid; Num/Den: Numerator/Denominator.

Source: Reviewer's analysis based on ADEFF.xpt dataset located at

[\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

Table 38: Proportion of Subjects with no IRF and no SRF in Central Subfield at Week 48, Observed cases, FAS (PULSAR)

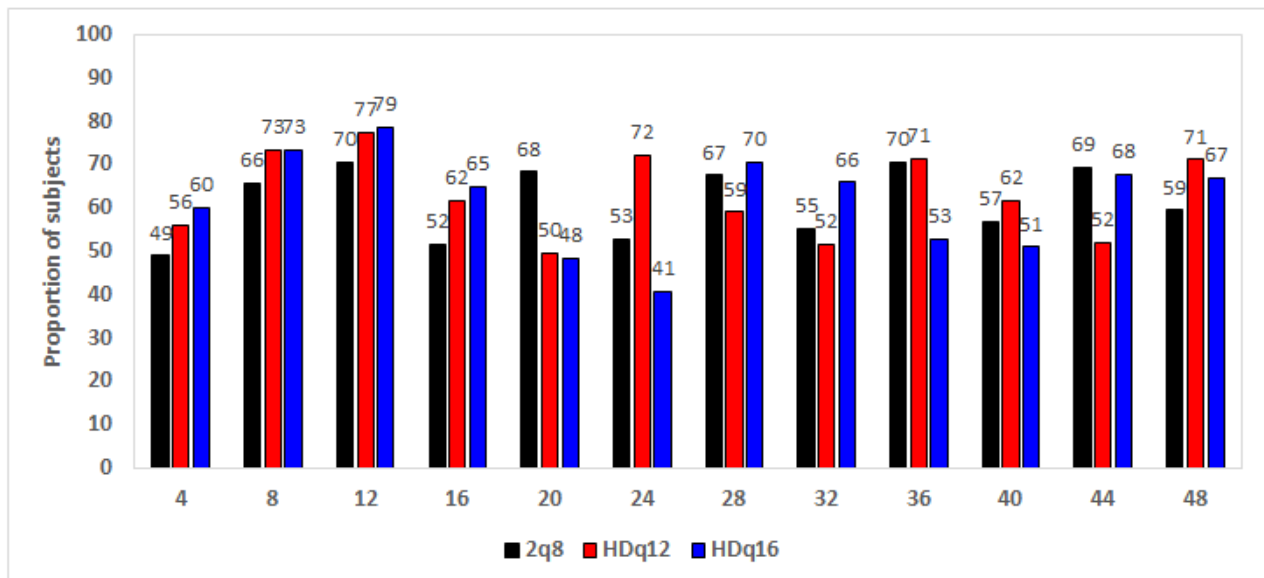
	2q8	HDq12	HDq16
Subjects who had no IRF and no SRF, Num/Den (%)	159/282 (56.4%)	214/298 (71.8%)	193/285 (67.7%)
Difference % (95% CI)		15.4 (7.7, 23.1)	11.4 (3.5, 19.3)

IRF: intraretinal fluid; SRF: subretinal fluid; Num/Den: Numerator/Denominator.

Source: Reviewer's analysis based on ADEFF.xpt dataset located at

[\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\20968\analysis\adam\datasets](#) for PULSAR.

Figure 18: Proportion of Subjects with no IRF and no SRF by Visit, LOCF (PULSAR)



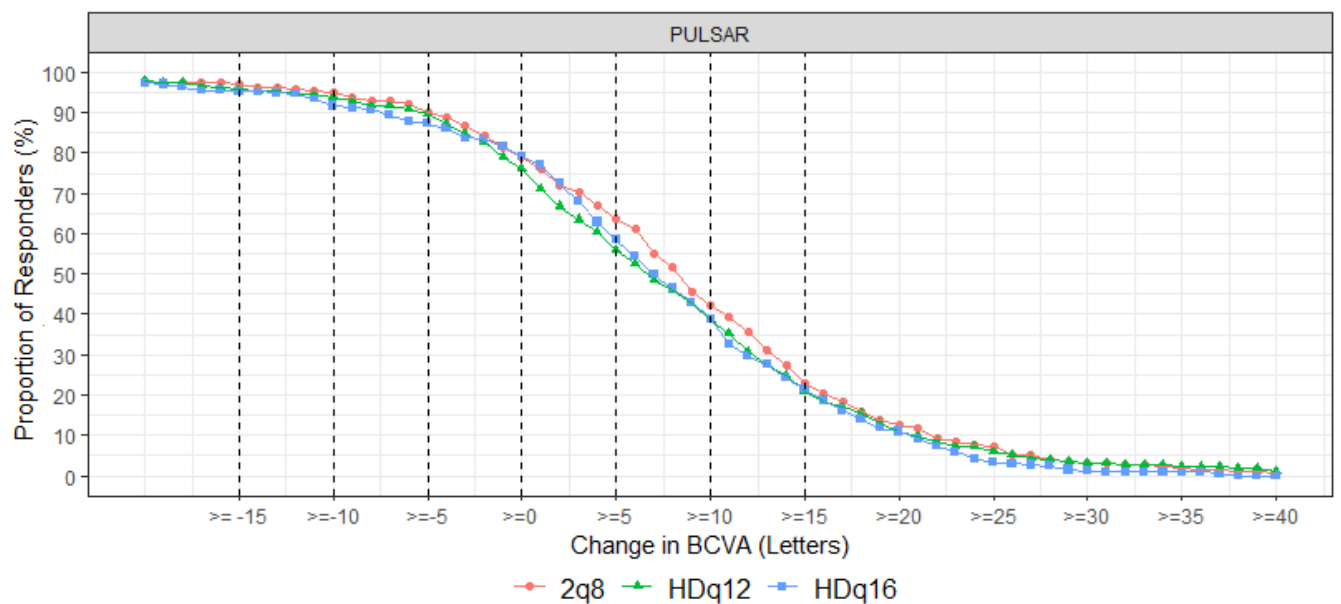
Source: Based on Reviewer's analysis

Table 39: Proportion of Subjects with no IRF and no SRF by Visit, LOCF (PULSAR)

Visit	2q8	HDq12	HDq16	HDq12 vs 2q8	HDq16 vs. 2q8
Week 4	164 (49.0%)	185 (56.1%)	199 (59.9%)	7.23 (-0.27, 14.73)	11.12 (3.65, 18.59)
Week 8	220 (65.7%)	244 (73.3%)	245 (73.4%)	7.80 (0.96, 14.64)	7.89 (1.00, 14.78)
Week 12	235 (70.4%)	258 (77.5%)	263 (78.5%)	7.36 (0.82, 13.90)	8.30 (1.78, 14.82)
Week 16	173 (51.6%)	205 (61.6%)	217 (65.0%)	10.04 (2.61, 17.47)	13.31 (5.98, 20.64)
Week 20	228 (68.3%)	165 (49.5%)	162 (48.4%)	-18.60 (-25.89, -11.30)	-19.85 (-27.11, -12.58)
Week 24	177 (52.8%)	239 (72.0%)	136 (40.7%)	19.43 (12.32, 26.54)	-12.01 (-19.48, -4.53)
Week 28	226 (67.5%)	196 (59.0%)	236 (70.4%)	-8.36 (-15.58, -1.13)	3.13 (-3.82, 10.07)
Week 32	185 (55.2%)	171 (51.5%)	220 (65.9%)	-3.65 (-11.20, 3.89)	10.76 (3.45, 18.06)
Week 36	236 (70.4%)	236 (71.3%)	177 (52.8%)	0.89 (-5.95, 7.73)	-17.56 (-24.77, -10.36)
Week 40	190 (56.9%)	204 (61.6%)	170 (50.9%)	4.68 (-2.75, 12.11)	-6.01 (-13.55, 1.53)
Week 44	232 (69.3%)	172 (51.8%)	227 (67.8%)	-17.45 (-24.72, -10.18)	-1.36 (-8.35, 5.63)
Week 48	199 (59.4%)	236 (71.1%)	223 (66.8%)	11.72 (4.57, 18.88)	7.45 (0.18, 14.72)

Source: Based on Reviewer's analysis

Figure 19: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (PULSAR).



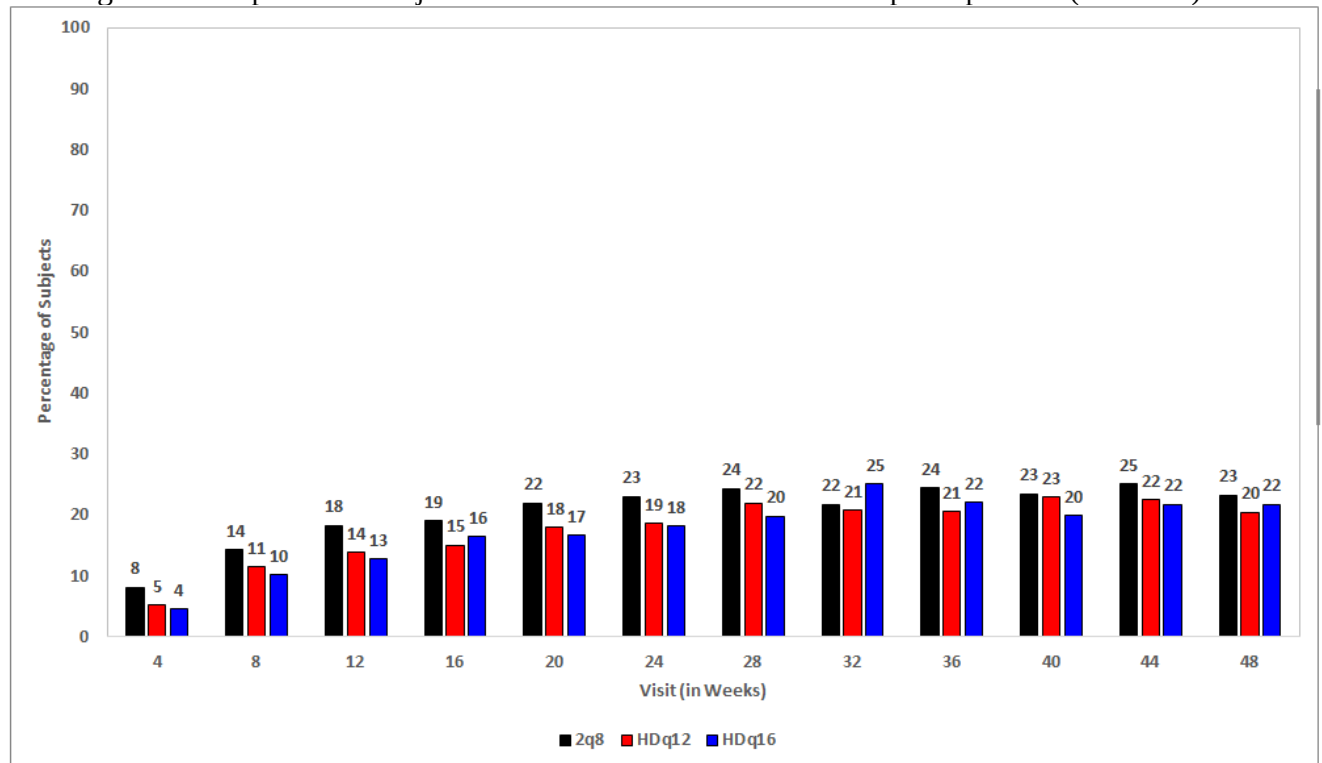
Source: Based on Reviewer's analysis

Table 40: BCVA-related Binary Efficacy Endpoints – Multiple Imputation (PULSAR)

Type	2q8	HDq12	HDq16	Adjusted Difference (95% CI)	
				HDq12 vs. 2q8	HDq16 vs. 2q8
Gained >= 15 Letters	23.2%	20.4%	21.7%	-3.1% (-9.4%, 3.2%)	-2.1% (-8.4%, 4.3%)
Gained >= 10 Letters	42.0%	38.9%	39.1%	-3.4% (-11%, 4.1%)	-3.3% (-11%, 4.2%)
Gained >= 5 Letters	63.1%	56.4%	58.3%	-7.0% (-15%, 0.6%)	-5.0% (-13%, 2.6%)
Gained >= 0 Letters	78.0%	76.4%	78.4%	-1.6% (-8.2%, 5.1%)	0.5% (-6.0%, 7.0%)
Avoided Losing >= 5 Letters	88.9%	89.1%	86.7%	-0.1% (-5.2%, 5.1%)	-2.7% (-8.1%, 2.7%)
Avoided Losing >= 10 Letters	94.2%	93.1%	91.3%	-1.4% (-5.4%, 2.6%)	-3.4% (-7.6%, 0.9%)
Avoided Losing >= 15 Letters	96.9%	95.5%	94.9%	-1.8% (-4.9%, 1.4%)	-3.1% (-6.4%, 0.1%)

Source: Based on Reviewer's analysis

Figure 20: Proportion of subjects who achieved >= 15 letters – Multiple Imputation (PULSAR)



Source: Based on Reviewer's analysis

Appendix 2: PHOTON Study

Table 41: Adjusted Mean Change in BCVA from Baseline Through Week 48 by Treatment (PHOTON)

	Difference in LS Means (95% CI)				
	2q8	HDq12	HDq16	HDq12 vs 2q8	HDq16 vs. 2q8
Week 4	4.8 (0.6)	4.2 (0.4)	3.8 (0.5)	-0.6 (-1.8, 0.6)	-1.0 (-2.3, 0.3)
Week 8	6.3 (0.6)	5.8 (0.5)	5.3 (0.6)	-0.6 (-1.9, 0.7)	-1.1 (-2.5, 0.4)
Week 12	6.7 (0.7)	6.5 (0.5)	5.6 (0.6)	-0.3 (-1.7, 1.2)	-1.1 (-2.7, 0.4)
Week 16	7.0 (0.7)	6.7 (0.5)	6.4 (0.6)	-0.3 (-1.9, 1.2)	-0.6 (-2.3, 1.1)
Week 20	7.6 (0.8)	6.1 (0.6)	6.0 (0.6)	-1.5 (-3.2, 0.2)	-1.6 (-3.3, 0.1)
Week 24	7.0 (0.8)	6.9 (0.6)	5.3 (0.7)	-0.1 (-1.8, 1.6)	-1.8 (-3.6, 0.1)
Week 28	7.8 (0.7)	7.3 (0.6)	6.8 (0.7)	-0.5 (-2.1, 1.1)	-1.0 (-2.8, 0.8)
Week 32	7.6 (0.8)	6.6 (0.6)	7.0 (0.7)	-1.0 (-2.7, 0.7)	-0.6 (-2.5, 1.2)
Week 36	8.3 (0.7)	7.4 (0.6)	6.2 (0.7)	-0.9 (-2.6, 0.8)	-2.1 (-4.0, -0.2)
Week 40	7.8 (0.7)	7.6 (0.6)	6.2 (0.7)	-0.2 (-1.9, 1.4)	-1.6 (-3.4, 0.2)
Week 44	8.5 (0.8)	7.7 (0.6)	7.1 (0.7)	-0.8 (-2.5, 0.9)	-1.4 (-3.3, 0.5)
Week 48	8.7 (0.7)	8.1 (0.6)	7.2 (0.7)	-0.6 (-2.3, 1.1)	-1.4 (-3.3, 0.4)

Source: Based on Reviewer's analysis

Table 42: Change from baseline in BCVA at Week 48 - Multiple Imputation with Tipping Point Analysis (PHOTON)

Shift value	LS Means (SE)			Difference in LS Means (95% CI)	
	2q8	HDq12	HDq16	HDq12 vs. 2q8	HDq16 vs. 2q8
0	8.4 (0.81)	7.9 (0.66)	6.9 (0.83)	-0.6 (-2.3, 1.1)	-1.5 (-3.4, 0.4)
-1	8.5 (0.82)	7.7 (0.66)	6.9 (0.83)	-0.7 (-2.4, 1.0)	-1.6 (-3.5, 0.4)
-2	8.5 (0.82)	7.6 (0.67)	6.8 (0.84)	-0.9 (-2.6, 0.8)	-1.7 (-3.6, 0.3)
-3	8.5 (0.83)	7.5 (0.67)	6.8 (0.84)	-1.0 (-2.8, 0.7)	-1.8 (-3.7, 0.2)
-4	8.6 (0.83)	7.4 (0.68)	6.7 (0.85)	-1.2 (-2.9, 0.5)	-1.8 (-3.8, 0.1)
-5	8.6 (0.84)	7.2 (0.68)	6.7 (0.85)	-1.4 (-3.1, 0.4)	-1.9 (-3.9, 0.1)
-6	8.6 (0.85)	7.1 (0.69)	6.6 (0.86)	-1.5 (-3.3, 0.3)	-2.0 (-4.0, 0.0)
-7	8.7 (0.85)	7.0 (0.69)	6.6 (0.87)	-1.7 (-3.4, 0.1)	-2.1 (-4.1, -0.1)
-8	8.7 (0.86)	6.9 (0.70)	6.5 (0.88)	-1.8 (-3.6, -0.0)	-2.2 (-4.2, -0.1)
-9	8.7 (0.87)	6.8 (0.71)	6.5 (0.89)	-2.0 (-3.8, -0.1)	-2.3 (-4.3, -0.2)
-10	8.8 (0.88)	6.6 (0.72)	6.4 (0.90)	-2.1 (-4.0, -0.3)	-2.3 (-4.5, -0.2)
-11	8.8 (0.89)	6.5 (0.72)	6.4 (0.91)	-2.3 (-4.1, -0.4)	-2.4 (-4.6, -0.3)
-12	8.8 (0.91)	6.4 (0.73)	6.3 (0.92)	-2.4 (-4.3, -0.5)	-2.5 (-4.7, -0.4)

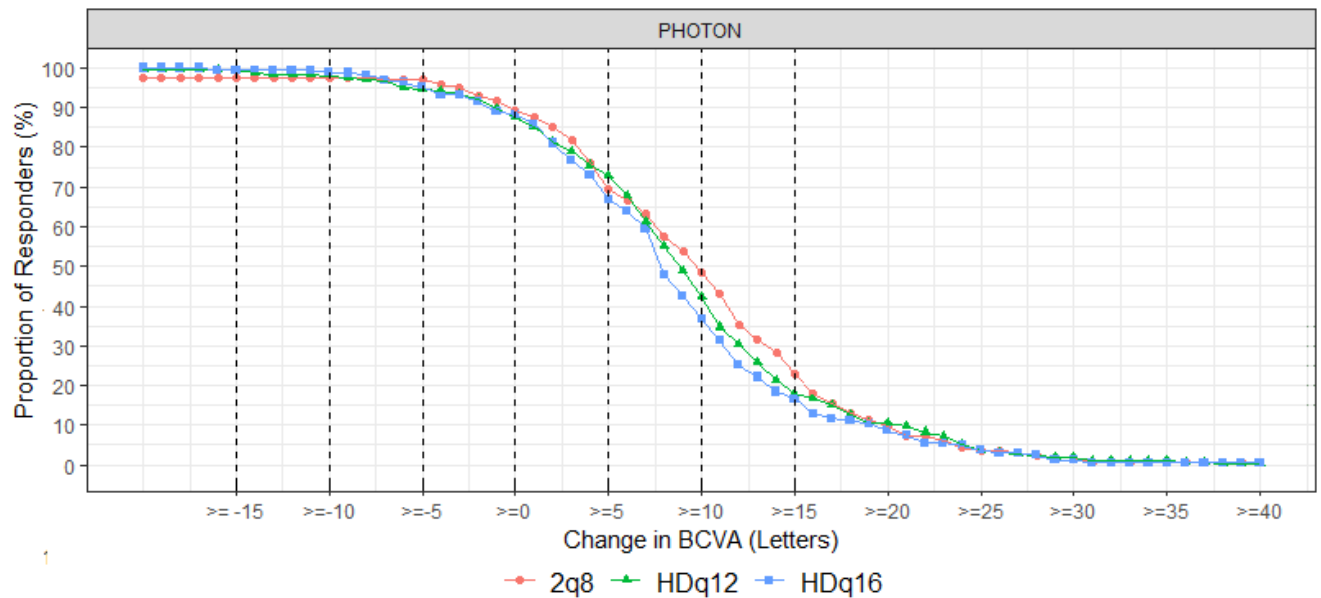
LS Means: Least square Means; SE: Standard error.

Shift value = 0 represents results based on the standard multiple imputation (MI) approach without a shift.

Shift value -1 to -12 represent Tipping Point analysis results.

Using a noninferiority (NI) margin of -4 letters, NI was established for shift value = 0. This conclusion of noninferiority was reversed at shift value -10 for HDq12 vs. 2q8 and at shift value -6 for HDq16 vs. 2q8.

Figure 21: Cumulative Distribution of the Change in BCVA from Baseline at Week 48 (PHOTON).



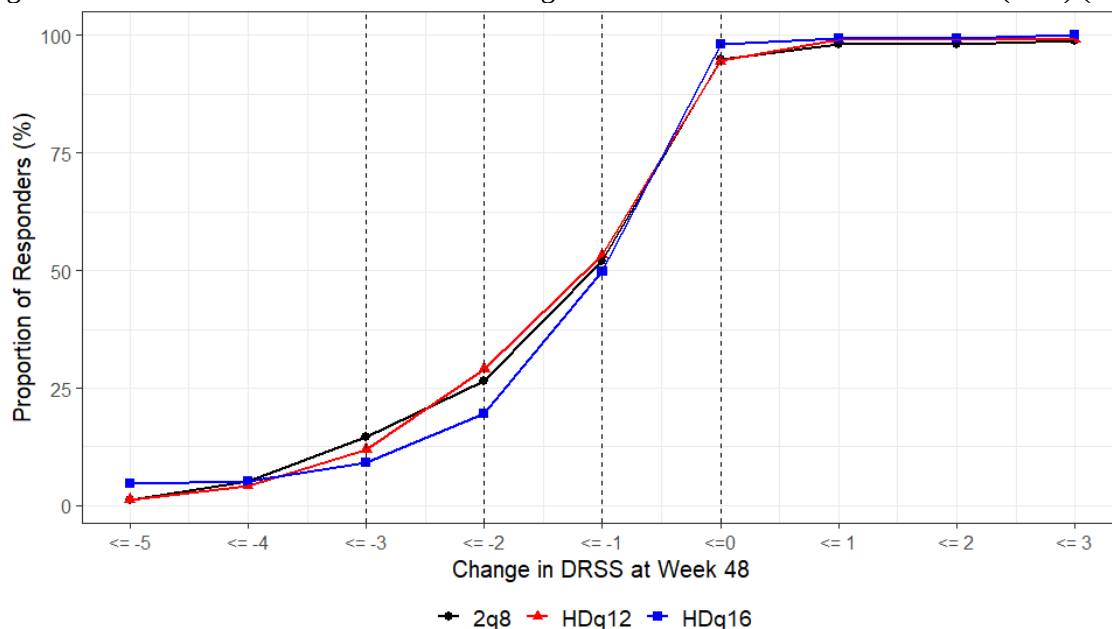
Source: Based on Reviewer's analysis

Table 43: BCVA-related Binary Efficacy Endpoints – Multiple Imputation (PHOTON)

Type	2q8	HDq12	HDq16	Adjusted Difference (95% CI)	
				HDq12 vs. 2q8	HDq16 vs. 2q8
Gained $\geq$ 15 Letters	24.0%	19.5%	17.5%	-4.8% (-13%, 3.0%)	-7.2% (-16%, 1.3%)
Gained $\geq$ 10 Letters	49.7%	42.4%	37.1%	-7.7% (-17%, 1.7%)	-13% (-24%, -3.0%)
Gained $\geq$ 5 Letters	69.5%	71.6%	66.2%	2.4% (-6.5%, 11.2%)	-4.2% (-15%, 6.2%)
Gained $\geq$ 0 Letters	89.4%	86.3%	87.5%	-4.0% (-10%, 2.3%)	-2.5% (-9.6%, 4.6%)
Avoided Losing $\geq$ 5 Letters	98.0%	94.0%	95.1%	0.1% (-6.7%, 7.0%)	-2.9% (-11%, 5.6%)
Avoided Losing $\geq$ 10 Letters	98.7%	97.5%	98.6%	NE	3.1% (-6.2%, 12.4%)
Avoided Losing $\geq$ 15 Letters	98.8%	98.9%	99.4%	NE	NE

Source: Based on Reviewer's analysis

Figure 22: Cumulative Distribution of the Change in DRSS from Baseline at Week 48 (FAS) (PHOTON)



Source: Based on reviewer's analysis

Table 44: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline to Week 48 – Multiple Imputation (Full Analysis Set)

Summary	2q8 (N = 158)	HDq12 (N = 310)	HDq16 (N = 153)
Achieved ≥ 2 -step improvement	26.2%	30.8%	20.3%
Adjusted Difference (95% CI)	---	4.0% (-5.0%, 13.1%)	-6.6% (-16.4%, 3.2%)
Achieved ≥ 3 -step improvement	14.9%	13.2%	9.5%
Adjusted Difference (95% CI)	---	-1.8% (-8.8%, 5.2%)	-5.5% (-13.1%, 2.1%)

Source: Based on reviewer's analysis

Table 45: Proportion of Subjects who Achieved ≥ 2 -Step and ≥ 3 -Step Improvement in DRSS from Baseline to Week 48 – Observed Cases (Full Analysis Set)

Summary	2q8 (N = 127)	HDq12 (N = 250)	HDq16 (N = 133)
Achieved ≥ 2 -step improvement	33/ 127 (26.0%)	77/ 250 (30.8%)	26/ 133 (19.5%)
Adjusted Difference (95% CI)	---	4.2% (-5.2%, 13.5%)	-7.4 (-17.5, 2.7)
Achieved ≥ 3 -step improvement	20/ 127 (15.7%)	34/ 250 (13.6%)	13/ 133 (9.8%)
Adjusted Difference (95% CI)	---	-2.3% (-9.9%, 5.3%)	-6.4% (-14.4%, 1.6%)

Source: Based on reviewer's analysis

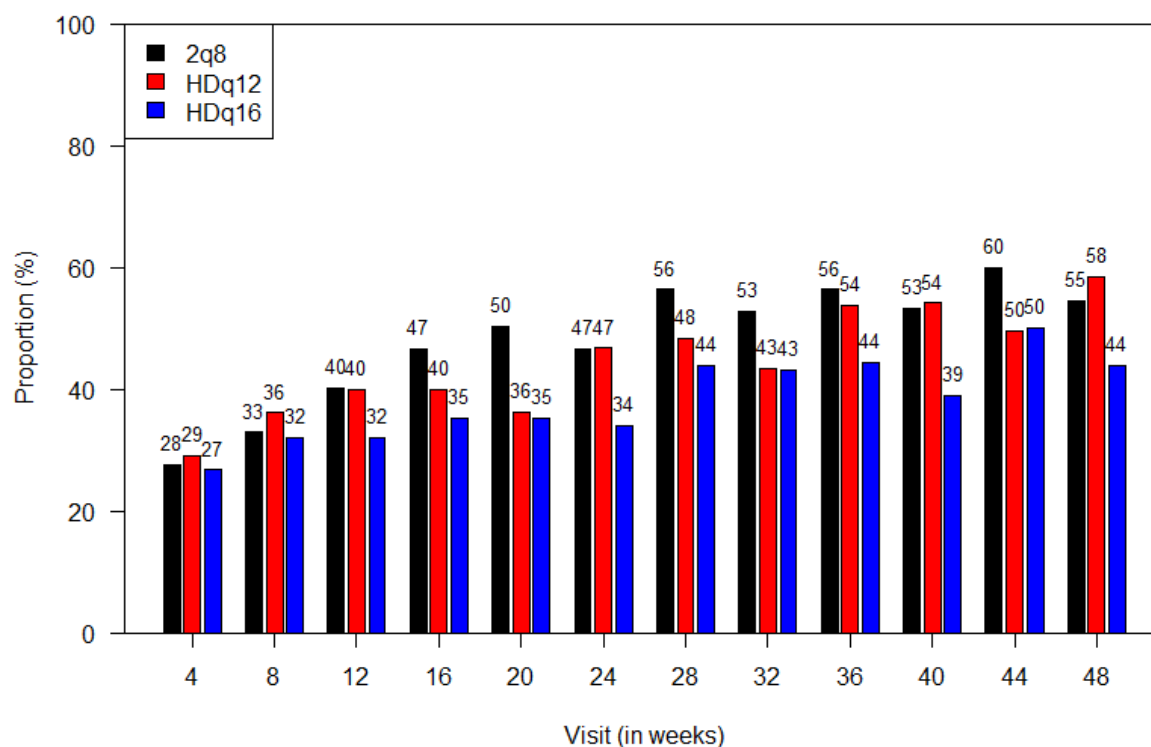
Table 46: Proportion of Subjects with no Fluid at Foveal Center at Week 48, OC (Full Analysis Set)

	2q8	HDq12	HDq16
Subjects who had no fluid, Num/Den (%)	77/148 (52.0%)	170/276 (61.6%)	66/149 (44.3%)
Difference % (95% CI)	--	9.5 (-0.2, 19.2)	-7.4 (-18.5, 3.6)

OC: Observed cases; Num/Den: Numerator/Denominator.

Source: Reviewer's analysis based on ADMOOCT.xpt dataset located at [\\CDSESUB1\evsprod\BLA761355\0001\m5\datasets\vgtfthd-dme-1934\analysis\adam\datasets](#) for PHOTON.

Figure 23: Proportion of Subjects with No Fluid in the Foveal Center by Visit, LOCF



Source: Based on Reviewer's analysis

Table 47: Proportion of Subjects with No Fluid in the Foveal Center by Visit, LOCF

Visit	2q8	HDq12	HDq16	HDq12 vs 2q8	HDq16 vs. 2q8
Week 4	44 (27.7%)	94 (29.1%)	43 (26.7%)	1.9 (-6.5, 10.3)	-0.1 (-9.7, 9.4)
Week 8	54 (32.9%)	117 (36.1%)	52 (32.1%)	3.8 (-5.0, 12.5)	-0.3 (-10.3, 9.7)
Week 12	66 (40.2%)	130 (40.0%)	52 (32.1%)	0.2 (-8.8, 9.2)	-7.6 (-17.8, 2.6)
Week 16	77 (46.7%)	130 (40.0%)	57 (35.2%)	-6.0 (-15.1, 3.0)	-10.9 (-21.3, -0.6)
Week 20	83 (50.3%)	118 (36.3%)	57 (35.2%)	-13.6 (-22.7, -4.4)	-14.9 (-25.4, -4.4)
Week 24	77 (46.7%)	152 (46.8%)	55 (34.0%)	0.7 (-8.5, 9.8)	-12.4 (-22.8, -2.1)
Week 28	93 (56.4%)	157 (48.3%)	71 (43.8%)	-7.7 (-16.9, 1.4)	-12.2 (-22.8, -1.6)

Visit	2q8	HDq12	HDq16	HDq12 vs 2q8	HDq16 vs. 2q8
Week 32	87 (52.7%)	141 (43.4%)	70 (43.2%)	-9.2 (-18.3, -0.0)	-9.2 (-19.8, 1.4)
Week 36	93 (56.4%)	175 (53.8%)	72 (44.4%)	-2.3 (-11.5, 6.8)	-11.7 (-22.3, -1.0)
Week 40	88 (53.3%)	176 (54.2%)	63 (38.9%)	1.0 (-8.1, 10.1)	-14.1 (-24.6, -3.5)
Week 44	99 (60.0%)	161 (49.5%)	81 (50.0%)	-10.2 (-19.1, -1.2)	-9.2 (-19.8, 1.3)
Week 48	90 (54.5%)	190 (58.5%)	71 (43.8%)	4.3 (-4.7, 13.3)	-9.7 (-20.3, 0.8)

Source: Based on Reviewer's analysis

Table 48: DRM Criteria in PULSAR and PHOTON Studies During Year 1

	PULSAR	PHOTON
Shorten dosing interval (Year 1)	BCVA loss >5 letters from Week 12, AND >25 µm increase in CRT from Week 12 OR new foveal hemorrhage OR new foveal neovascularization	BCVA loss >10 letters from Week 12 in association with persistent or worsening DME AND >50 µm increase in CRT from Week 12

BCVA = best corrected visual acuity, CRT = central retinal thickness, DRM = dose regimen modification

Source: Table 1 from the Summary of Clinical Efficacy.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

JUAN C VIVAR
05/26/2023 10:52:52 AM

SOLOMON CHEFO
05/26/2023 10:57:58 AM