Document Type:	Statistical Analysis Plan
Official Title:	An open-label, randomized, active-controlled, parallel-group, Phase-3b study of the efficacy, safety, and tolerability of 2 mg aflibercept administered by intravitreal injections using two different treatment regimens to subjects with neovascular agerelated macular degeneration (nAMD)
NCT Number:	NCT02540954
Document Date:	31-Jan-2020

An open-label, randomized, active-controlled, parallel-group, Phase-3b study of the efficacy, safety, and tolerability of 2 mg aflibercept administered by intravitreal injections using two different treatment regimens to subjects with neovascular age-related macular degeneration (nAMD)

Efficacy and safety of two different aflibercept regimens in subjects with nAMD

Bayer study drug	BAY86-5321/ aflibercept / VEGF Trap-Eye (Eylea)		
Study purpose:	Post-approval commitment		
Clinical study phase:	3b	**Date:**	31 Jan 2020
Study No.:	BAY86-5321 / 16598	**Version:**	3.0
Author:	PPD Syneos Health 81929 Munich Germany		
Syneos Health Project Code:	04.6000.1005563		

Confidential

The information provided in this document is strictly confidential and is intended solely for the guidance of the clinical investigation. Reproduction or disclosure of this document, whether in part or in full, to parties not associated with the clinical investigation or its use for any other purpose without the prior written consent of the sponsor is not permitted.

Throughout this document, symbols indicating proprietary names (®, TM) are not displayed. Hence, the appearance of product names without these symbols does not imply that these names are not protected.

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

Abbreviations

2Q8	2 mg aflibercept administered every 8 weeks
AE	adverse event
ALT	alanine aminotransferase
AMD	age-related macular degeneration
ANCOVA	analysis of covariance
APTC	Antiplatelet Trialists' Collaboration
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
ATE	arterial thrombotic event
BCVA	best corrected visual acuity
BHC	Bayer HealthCare
BL	baseline
BMI	body mass index
BRVO	branch retinal vein occlusion
BUN	blood urea nitrogen
CI	confidence interval
CNV	choroidal neovascularization
CRF	case record form
CRT	central retinal thickness
CRVO	central retinal vein occlusion
DME	diabetic macular edema
DMP	data monitoring plan
eCRF	electronic case report form
EMA	European Medicines Agency
ETDRS	Early Treatment Diabetic Retinopathy Study
EU	European Union
EudraCT	EU Drug Regulating Authorities Clinical Trials
FA	fluorescein angiography
FAS	full analysis set
FP	fundus photography
FPFV	First-Patient-First-Visit
HDL	high-density lipoprotein
INR	international normalized ratio
IOP	intraocular pressure

IVT	intravitreal(ly)
LDM	Lead Data Manager
LOCF	last observation carried forward
LS	least squares
MI	Multiple Imputations
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
mCNV	myopic choroidal neovascularization
MCV	mean corpuscular volume
MedDRA	Medical Dictionary for Regulatory Activities
nAMD	Neovascular macular degeneration
NEI VFQ-25	National Eye Institute 25-item Visual Function Questionnaire
OC	observed cases
OCT	optical coherence tomography
PDD	protocol deviation document
PPS	per-protocol set
PT/INR	prothrombin time
PT	preferred term
PTT	partial thromboplastin time
RMM	repeated measurements model
SAE	serious adverse event
SAF	safety population
SAP	statistical analysis plan
SOC	system organ class
SOP	standard operating procedure
TEAE	treatment-emergent AE
UPCR	urine protein/creatinine ratio
US	United States
VEGF	vascular endothelial growth factor
VIEW 1	Vascular Endothelial Growth Factor VEGF Trap-Eye: Investigation of Efficacy and Safety in Wet Age-Related Macular Degeneration (AMD)
VIEW 2	A randomized, double masked, active controlled, phase 3 study of the efficacy, safety, and tolerability of repeated doses of intravitreal VEGF Trap-Eye in subjects with neovascular age-related macular degeneration (AMD)
VRM	Validity Review Meeting
VTE	VEGF Trap-Eye
WHODD	World Health Organization Drug Dictionary

1 Introduction

1.1 Objective

This statistical analysis plan (SAP) describes the statistical methods and data presentations to be used in the summary and analyses of data from an open-label, randomized, active-controlled, parallel-group, Phase-3b study of the efficacy, safety, and tolerability of 2 mg aflibercept administered by intravitreal injections using two different treatment regimens to subjects with neovascular age-related macular degeneration (nAMD).

This document is based on the Clinical Study Protocol No. BAY 86-5321 /16598, version 2.0, dated 10 December 2015. It describes the plans for the final analyses.

1.2 Background

Age-related macular degeneration (AMD) is a leading cause of adult blindness in the developed world (5). Age-related macular degeneration has a dry and a wet form, the latter of which accounts for most AMD-related cases of blindness and is referred to as neovascular AMD (nAMD). Severe vision loss from nAMD is caused by a combination of retinal edema and neovascular proliferation. Vascular endothelial growth factor (VEGF), a protein growth factor that both stimulates angiogenesis and increases vascular permeability, is a major pathogenic factor in AMD (1). Anti-VEGF therapy has been shown to provide significant therapeutic benefit to patients suffering from nAMD.

Aflibercept is a potent, specific inhibitor of VEGF with a high affinity for all isoforms of VEGF and placental growth factor. To date, aflibercept has been approved as a treatment for nAMD in the United States (US), European Union (EU), Japan and several other countries and for the treatment of macular edema secondary to central retinal vein occlusion (CRVO) in the US and several other countries. Phase-3 studies of aflibercept in diabetic macular edema (DME), branch retinal vein occlusion (BRVO), and myopic choroidal neovascularization (mCNV) are ongoing.

Further details can be found in the latest available version of the investigator's brochure, which contains comprehensive information on the study drug.

The approval of aflibercept in the EU (22 November 2012) was based upon the results of two randomized, multicenter, double-masked, active-controlled Phase-3 clinical studies — the VIEW 1 (US and Canada) and VIEW 2 (EU, Asia/Pacific, and Latin America) studies. Both studies demonstrated after 52 weeks that aflibercept, dosed every 8 weeks following three initial monthly injections, provided efficacy that was non-inferior to monthly dosing of ranibizumab. The primary endpoint of these studies was the proportion of patients who maintained visual acuity (less than 15 letters of vision lost on the Early Treatment Diabetic

Retinopathy Study [ETDRS] chart) over 52 weeks. During the second year of these studies, the treatment interval could be extended (based on protocol-specified re-treatment criteria), with no more than 12 weeks allowed between injections, while the label allows an extension of the treatment interval beyond 12 weeks. To address the remaining issue regarding the optimal dosing frequency after 1 year of treatment with aflibercept, EMA requested a post-authorization randomized study with primary objective of comparing the proactive 8-weekly injections to a reactive regimen based on visual and anatomic outcomes.

The approved EU labeling for aflibercept at the time-point of study design stated:

> The recommended dose for Eylea is 2 mg aflibercept, equivalent to 50 microliters.
>
> Eylea treatment is initiated with one injection per month for three consecutive doses, followed by one injection every two months. There is no requirement for monitoring between injections.
>
> After the first 12 months of treatment with Eylea, the treatment interval may be extended based on visual and anatomic outcomes. In this case the schedule for monitoring should be determined by the treating physician and may be more frequent than the schedule of injections.

1.3 Rationale of the Study

As a condition for approval, the European Medicines Agency (EMA) has required a study to assess every-other-month dosing versus an extended-dosing regimen with no maximum limit to the treatment interval. This Phase-3b post-authorization efficacy study has been designed to meet this obligation and compares a fixed dosing regimen of 2 mg aflibercept administered every 8 weeks (“2Q8 group”) to a flexible extended regimen of 2 mg aflibercept (“extended dosing group”) in patients with nAMD who have completed at least 1 year of treatment with aflibercept.

1.4 List of Documents used

- Clinical Study Protocol No. BAY 86-5321 /16598, version 2.0, dated 10 December 2015
- Protocol deviation document, version 8.0, dated 06 November 2018

2 Study Objectives

Primary objective

To compare the efficacy of 2 mg aflibercept administered by two different intravitreal (IVT) treatment regimens to subjects with nAMD.

Secondary objective

To assess the safety and tolerability of aflibercept in this patient population

3 Study Design

3.1 Design Overview

This is a randomized, 2-arm, active-controlled, parallel-group, open-label, multicenter, Phase-3b study to assess the non-inferiority of an extended-dosing regimen to a 2Q8 fixed dosing regimen in subjects with nAMD who have completed at least 1 year of treatment with aflibercept (from first treatment to randomization into this study) in accordance with the constraints as provided in this protocol.

This study comprises a screening period of up to 66 days. There will be two parallel treatment arms. In both arms, aflibercept will be injected IVT at a dose of 2 mg per injection. The treatment arms differ in dosing intervals:

VTE extended dosing:	Flexible dosing intervals of ≥8 weeks (no upper limit) based on visual and anatomic outcomes as judged by the investigator. When/if visual and anatomical outcomes indicate that the disease has reactivated, the treatment interval will revert to the last treatment interval in which the disease was inactive (i.e. no signs of exudation were observed).
VTE 2Q8:	Fixed dosing intervals of 8 weeks (± 3 days) - modification of the treatment interval is not allowed. VTE 2Q8 will serve as the control-group.

3.2 Inclusion/Exclusion

Two sets of inclusion and exclusion criteria will be used to select subjects for participation in this study:

The first set of inclusion/exclusion criteria will be based on the subject's medical history prior to initiation of treatment with marketed aflibercept and will be assessed by a review of the subject's medical records. This will ensure that the subject's pre-aflibercept status matches, as close as possible, the baseline (BL) criteria of the pivotal VIEW program.

- 1st set of inclusion-criteria: incl. 1 to incl. 3
- 1st set of exclusion-criteria: excl. 1 to excl. 7

The second set of inclusion/exclusion criteria considers both the subject's medical history (assessed by a review of the subject's medical records) and the actual status of the subject at the time of signing the informed consent form (i.e. at enrollment into the study via conduct of study-specific screening procedures).

- 2nd set of inclusion-criteria: incl. 4 to incl. 8
- 2nd set of exclusion-criteria: excl. 8 to excl. 37

Details are available in the clinical study protocol, Section 6.1 and 6.2.

3.3 Visit Overview

Following screening, eligible subjects will be randomized (1:1) to one of the two parallel treatment group. Subjects in both treatment groups will receive the first study treatment on Day 1. The treatment period in the 2Q8 group is 72 weeks with a final study visit at Week 76, the treatment period in the VTE extended dosing group is 76 weeks.

Two post-baseline visits will follow the same strict schedule in both treatment groups: primary completion at Week 52 and the final visit at Week 76. All other post-baseline visits may differ between the treatment groups as detailed in study-protocol, Section 9.

The last possible study drug injection in the extended-dosing group can be given at Week 76 (i.e. the end of the study) after completion of all protocol-specified assessments[1]. After Week 76, subjects should be returned to standard of care.

For each treatment group, the time window for all post-baseline visits is ± 3 days relative to baseline

Table 1 gives a description of the planned assessments and study procedures.

3.4 Efficacy and Safety variables

The primary efficacy variable will be the change in best corrected visual acuity (BCVA) using ETDRS chart letter score for the study eye (starting at 4 meters) from baseline to Week 52. A complete list of all efficacy variables is provided in Section 6.1.6.

Both the primary and secondary efficacy variables will be assessed at Week 52/last observation carried forward (LOCF). The primary analysis will be based on the Full Analysis Set.

Efficacy will be assessed by the change in ETDRS BCVA letter score from study baseline. BCVA will be evaluated at every study visit. Other measures of efficacy will include changes from baseline in central retinal thickness (CRT) and choroidal neovascularization (CNV) as measured by optical coherence tomography (OCT) at each visit, and changes from baseline in

[1] If an extended-dosing subject receives an injection at Week 76, it is the responsibility of the treating investigator to follow-up on any adverse events (including ongoing events) that may occur within 4 weeks following this treatment. Information regarding such events is to be reported under this protocol (i.e. not as spontaneous reports).

quality of life as assessed by the National Eye Institute 25-item Visual Function Questionnaire (NEI VFQ-25) at baseline, Week 52 and Week 76. Proportion of subjects with extended treatment interval and number of injections are also considered as efficacy measures,

Assessments of ocular safety will include intraocular pressure (IOP), indirect ophthalmoscopy, and slit lamp biomicroscopy at all study visits. Under this protocol, mandatory fluorescein angiography (FA)/fundus photography (FP) examinations will be conducted during screening and at Weeks 52 and 76 only. However, the treating investigator may perform FA/FP at other times during the study based on his/her medical judgment and standard of care.

Overall safety of the subjects will be assessed throughout the study by monitoring ophthalmic and systemic adverse events (AEs). All potential arterial thrombotic events (ATEs) will be adjudicated according to the Antiplatelet Trialists' Collaboration (APTC) endpoints of non-fatal myocardial infarction, non-fatal stroke, and fatal vascular events. Vital signs will be assessed at all study visits and laboratory tests (hematology parameters, blood chemistry parameters, and parameters of urinalysis) will be performed at study baseline and at Week 76 only.

Table 1: Schedule of assessments and study procedures

Visit number [a]	*Screening* Visit 1	*Baseline (BL)* Visit 2	Visits 3 - 8	*Prim. compl.* Visit 9	Visits 10 - 12	*Final visit* Visit 13[h]
Time point[a] 2Q8 group	up to 66 days prior to BL (both groups)	Day 1 (both groups)	Weeks 8 – 48	Week 52 (both groups)	Week 56, 64, 72	Week 76 (both groups)
Extended-dosing group			Flex. intervals (≥ 8 weeks)		Flex. intervals (≥ 8 weeks)	
Initiation procedures						
Informed consent	●					
Demographic data	●					
Medical / ophthalmic history	●					
Physical examination	●					
PT / INR and PTT	●					
Inclusion / exclusion criteria	●	●				
Study medication						
Randomization		●				
Administration of study treatment [f]		●	●	extended-dosing only [i]	●	extended-dosing only [i,j]
Ophthalmologic assessments						
BCVA (ETDRS chart starting at 4 m) [g]	●	●	●	●	●	●
Optical coherence tomography	●	●	●	●	●	●
Fluorescein angiogr., fundus photogr.	●			●		●
Indirect ophthalmoscopy [d]	●	●	●	●	●	●
Slit lamp biomicroscopy	●	●	●	●	●	●
Intraocular pressure (IOP) [e]	●	●	●	●	●	●
Standard safety						
Prior / concomitant medications	●	●	●	●	●	●
Adverse events [c]		●	●	●	●	●
Vital signs (temp., blood pressure, HR)	●	●	●	●	●	●
Hematology / chemistry	●					●
Urinalysis / UPCR	●					●
Serum pregnancy test (women of childbearing potential only)	●					
Urine dipstick pregnancy test (women of childbearing potential only)		●				●
Other						
NEI VFQ-25 [b]		●		●		●

BCVA = Best corrected visual acuity; ETDRS = Early Treatment Diabetic Retinopathy Study; HR = heart rate; NEI VFQ-25 = National Eye Institute 25-item Visual Function Questionnaire; PT/INR = Prothrombin time / International normalized ratio; PTT = Partial thromboplastin time; UPCR = Urine protein/creatinine ratio

a: Visit schedules may deviate by ± 3 days relative to baseline. Scheduled visits should not be altered due to the deviation of the previous visit.
Visit numbers refer to the mandatory visits for treatment administration to subjects in the 2Q8 group. Fewer visits may be required for subjects in the extended-dosing group, for whom extended-dosing intervals are indicated.

b: The NEI VFQ-25 is to be administered in a quiet room by a person certified to administer the questionnaire

c: Any AE occurring up to 4 weeks after the last injection of aflibercept has to be documented, regardless of the relationship to the study drug or the seriousness of the event and reported in accordance with this protocol (i.e. not as a spontaneous report). For any drug-related AE occurring after 4 weeks after the last application of aflibercept, the standard procedures that are in place for spontaneous reporting will be followed.

d: Pre- and post-dose at visits with study drug administration

e: Pre-dose and 30-60 minutes post-injection at visits with study drug administration

f: See Section 16.1 of the study protocol for an example study drug injection protocol.
g: Refraction to be done at each visit
h: Also to be conducted in case of premature discontinuation of study participation
i: Extended-dosing subjects may receive injections at Week 52 or 76 depending on their individual schedule.
j: If an extended-dosing subject receives an injection at Week 76, follow-up is needed on any adverse events (including ongoing events) that may occur within 4 weeks following this treatment (AE reporting under this protocol; i.e. not as spontaneous reports)

4 General Statistical Considerations

The statistical evaluation will be performed by using the software package SAS release 9.4 or higher (SAS Institute Inc., Cary, NC, USA).

The statistical analysis plan (SAP) version 1.0 of this open-label study is finalized before First-Patient-First-Visit (FPFV) to avoid additional reporting bias.

4.1 General Principles

The variables will be analyzed by descriptive statistical methods. The number of data available and missing data, mean, standard deviation, minimum, quartiles, median, and maximum will be calculated for metric data. Mean and standard deviation will be reported to one decimal place greater than the data were collected. Quartiles, median, minimum and maximum values will be reported with the same precision as they were collected.

Frequency tables will be generated for categorical data. These include the counts and percentages of each category including the category ‘missing’ as a separate category, if applicable. Percentages will be calculated using a denominator of all subjects in the specified population and treatment-group, and the percentage values will be reported to one decimal place.

If not stated otherwise, ANCOVA-tables will include the least squares (LS) means at baseline, the LS mean changes relative to BL for both treatment groups, and the difference in the LS means (extended-dosing group minus 2Q8 group) as point estimate and as 95%-CI.

4.2 Handling of Dropouts

Subjects must or might be withdrawn from study for different reasons, which are specified in the protocol in Section 6.3.1. Subjects who withdraw from the study will not be replaced. Premature permanent discontinuation from study medication implies premature discontinuation from study participation.

Depending on the time point of withdrawal, a withdrawn subject is referred to as either “screening failure” or “dropout” as specified below:

Screening failure

A subject who, for any reason (e.g. failure to satisfy the selection criteria), terminates the study before the time point used for the definition of "dropout" (see below) is regarded a "screening failure".

Re-screening of screening failures may be acceptable under the specific conditions described in Section 6.3.1 of the study-protocol.

Drop-out

A subject who discontinues study participation prematurely for any reason is defined as a "drop-out" if the subject has already been randomized.

The number of screening failures and drop-outs as well as the respective reasons will be summarized.

4.3 Handling of Missing Data

All missing or incomplete data will be presented in the subject data listings as they are recorded on the Case Report Form (CRF). In general they will not be substituted or replaced, except for the parameters described below and in Section 6.1.6 for efficacy analysis.

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

4.3.1 Adverse Events

In general, data will not be imputed for safety analysis. If dates of adverse experiences (clinical or laboratory untoward events) are missing so that the determination of whether or not the event is treatment-emergent is questionable, the event will be presumed to be treatment-emergent. If it is clear from the (partially missing date) that the adverse event occurred prior to first dose of study drug or more than 30 days after the last dose of study drug, it will be treated as not treatment emergent.

4.3.2 Efficacy Analysis

The primary method for replacing missing values for all efficacy analyses will be "last observation carried forward" (LOCF), see Section 4.3.2.1.

The following other methods will be applied to account for missing data for the primary endpoint (assuming that the missing values are missing at random):

- Multiple imputations (MI), see Section 4.3.2.2.
- Repeated measurements model (RMM), see Section 4.3.2.3.

Moreover, an observed case analysis (OC), see Section 4.3.2.4, will be provided.

4.3.2.1 Last Observation Carried Forward

Missing ETDRS BCVA letter scores will be imputed by LOCF using the last available scheduled post-baseline values.

After the imputation the respective dataset will be filled and the data will be analyzed as if the missing values had been observed (as the last available value). For example, the primary efficacy variable will be analyzed as described in Section **Error! Reference source not found.**.

4.3.2.2 Multiple Imputation

Multiple imputations (MI) for missing values will additionally performed on the primary and key secondary efficacy variables.

Multiple imputation methods involve three steps:

I. Imputation
i.e., the generation of multiple copies of the original dataset by replacing missing values using an appropriate stochastic model.

a. First missing data will be imputed in order to achieve a monotone missing pattern using the MCMC (Markov Chain Monte Carlo) method, using SAS-procedure proc MI similarly as below.

```
PROC MI DATA=<indata> SEED=3818 OUT=out1 NIMPUTE=20;
      MCMC impute=monotone;
      VAR base t1 t2 … ;
RUN;
```

b. Subsequently missing data will be imputed by a regression model.

```
PROC MI DATA=out1 SEED=4144 OUT=full nimpute=1;
      BY _Imputation_;
      CLASS treatment;
      MONOTONE method=reg;
      VAR treatment base t1 t2 … ;
RUN;
```

II. Analysis,
i.e., the analysis of the multiple imputed datasets as complete sets. The analysis step is performed for each of the multiply imputed datasets. Since all imputed datasets are complete there is no need to bother with any missing data. On each imputed dataset the primary and key secondary analysis will be performed as described in Sections 0 and 6.2.2. It should be clarified, that classifications (loss of < 15 letters) will be performed after the imputation of the missing BCV-values.

III. Pooling,
i.e., the combination of the different parameter estimates across the multiple datasets based on Rubin's rules to produce a unique point estimate and standard error taking into account the uncertainty of the imputation process using SAS procedure proc MIANALYZE.

4.3.2.3 Repeated Measurements Model

The repeated measurements model (RMM) does not employ formal imputation. RMM has been extensively used in the analysis of longitudinal data especially when missing data is a concern and missing at random (MAR) is assumed.

For RMM, if SAS mixed model is used, the sample SAS codes will be like the following:

```
PROC MIXED DATA=<indata>;
     CLASS subject treatment(ref=VTE2Q8) time-window;
     MODEL chg = base treatment time-window treatment*time-
     window /ddfm=kr;
     repeated time-window / sub = subject type = un;
     lsmeans treatment / cl diff pdiff=CONTROL(2Q8fix);
     RUN;
```

Where the treatment difference is obtained with the lsmeans statement for the treatment differences at time Week52.

4.3.2.4 Observed Cases

The primary analysis will be repeated on observed cases, without any imputation.

4.3.3 Prior and Concomitant Medication / Medical History

Completely missing start and stop dates of medication are considered missing and no replacement is generated. A medication with a complete missing start-date will be assumed to start before first application of the study-drug. A complete missing stop-date will be handled as "ongoing".

The following algorithm will be used for the calculation of durations and/or to flag medications:

- for incomplete start dates: if only day is missing - first of month, if day and month missing - first day (1st January) of year.

- for incomplete stop dates: if only day is missing - last of month, if day and month missing - last day (31st December) of year or "ongoing", when last day of year is after last study-day of the subject.

4.4 Interim Analyses and Data Monitoring

No formal interim analysis will be conducted.

4.5 Data Rules

Determination of baseline values

Generally, pre-treatment values recorded at Visit 2 (Day 1) will be used as baseline values. This visit should take place within 8 weeks of the screening visit. If no baseline value is available then last available screening values are used. Change from baseline is calculated as the value at the post-baseline time point minus the baseline value, i.e. value at time point – value at baseline.

Handling of repeated measurements at the same visit

If measurements were repeated at the same scheduled visit, the value actually flagged as scheduled will be the

- Last non-missing repeated measurement, if measurement is before start of treatment, and
- First non-missing repeated measurement, if measurement is after start of treatment.

Generally, only scheduled measurements will be used for statistical summaries and analysis. Unscheduled measurements will not be used for analysis, however they will be listed.

Handling of scheduled time-windows

The screening visit must occur within 8 weeks of the baseline visit (Day 1). Screening values collected more than 8 weeks before baseline will be flagged in the patient listings, but might be used for summary tables and analysis.

For the scheduling of Visit 3 to Visit 13 (Week 4 to Week 76) ± 3 days relative to baseline are foreseen.

Values outside these windows will not be dropped from summary tables and analysis, but flagged as “outside time-window” in the patient listings.

Defining time-windows for the visit wise summaries

The visits for the control-group (VTE 2Q8) are fixed, therefore the respective results will be summarized as scheduled – i.e. Screening, Baseline, Week 8, Week 16, Week 24, Week 32, Week 40, Week 48, Week 52, Week 56, Week 64, Week 72 and Week 76, if applicable.

For the VTE extended dosing group the following time-windows after BL will be defined:

- Screening: study-days - 66 to -1
- Baseline: study-day 1
- Week 1-8: study-days 2 to 59
- Week 9-16: study-days 60 to 116

- Week 17-24: study-days 117 to 172
- Week 25-32: study-days 173 to 228
- Week 33-40: study-days 229 to 284
- Week 41-48: study-days 285 to 340
- Week 49-51: study-days 341 to 361
- Week 52: study-days 362 to 368
- Week 53-60: study-days 369 to 424
- Week 61-68: study-days 425 to 480
- Week 69-75: study-days 481 to 529
- Week 76: study days 530 to 536
- > Week 76: study days > 536

The last measurement before or at the upper bound of the time interval will be displayed in each time interval.

These rules especially apply for all assessments at End of Study / End of treatment in both treatment-groups.

Pooling centers

Two types of pooling will be performed:

a) All centers combined

b) Geographic region

Calculation of durations and study-days

Durations and study-days are calculated relative to baseline, if not specified otherwise.

Coding

The verbatim of the following panels will be coded by the latest version of Medical Dictionary for Regulatory Activities (MedDRA) available before database lock

- Medical history
- Adverse Events
- Surgeries after start of study

Prior and concomitant medications will be coded by the latest version of World Health Organization Drug classification Dictionary (WHO-DD) available before database lock.

Presentation

Listings will be sorted by treatment, unique subject identifier and date if applicable.

Dates will be formatted as DDMMMYYYY. Partial dates will be presented on data listings as recorded on CRFs.

Rounding for all variables will occur only as the last step, immediately prior to presentation in listings and tables. No intermediate rounding will be performed on derived variables. The standard rounding practice of rounding numbers ending in 0-4 down and numbers ending 5-9 up will be employed.

Every table, listing and figure will be produced with an electronic date stamp to document when it was produced.

4.6 Validity Review

Validity Review Meetings (VRMs) are performed according to Bayer HealthCare (BHC) Standard Operating Procedures (SOP) and will be led by the Syneos Health Lead Data Manager (LDM). Details are available in the DMP.

The results of the VRM will be documented in the Validity Review Report and may comprise decisions and details relevant for statistical evaluation. Any changes to the statistical analysis prompted by the results of the validity review meeting will be documented in an amendment and, if applicable, in a supplement to this SAP.

5 Analysis Sets and Subgroups

5.1 Analysis Sets

Populations for analysis will be defined as follows:

Full analysis set (FAS)

The FAS will include all randomized subjects who received any study drug and have a baseline BCVA assessment and at least one post-baseline BCVA assessment.

The FAS will be analysed as randomized.

Per-protocol set (PPS)

The PPS will include all FAS subjects who have at least one BCVA assessment at Week 36 or later and do not have a major protocol deviation. Major protocol deviations will be defined in the protocol deviation document (PDD). In version 8.0 of the PDD from November 6th, 2018 the following major protocol deviations are listed for FAS-subjects:

- In-/Exclusion criteria not met but subject entered treatment phase
- Excluded concomitant medication treatment
- Procedure deviations
    - Visit 2 has not been performed

- Planned BCVA-examination in study eye has not been performed at any Post Baseline Visit >= Week 36
- Patient missed two or more consecutive study drug injections prior to week 52

PPS will be defined at the last VRM i.e. prior to database lock. The PPS will be analysed as randomized.

Safety analysis set (SAF)

The SAF will include all subjects who receive any study drug under this protocol.

The safety analysis set will be analysed as randomized. Subjects treated other than randomized will be marked in the listings.

5.2 Definition of Subgroups

Subgroups are

- Sex (f/m)
- Geographic region
  - Canada and Western Europe: Canada , France and UK
  - Middle Europe: Austria, Germany and Switzerland
  - Eastern Europe: Czech Republic, Hungary, Poland, Slovakia and Lithuania
  - Southern Europe: Spain, Portugal and Italy

- Change of BCVA from first aflibercept administration to AZURE baseline (< 10 letters, >= 10 letters)

  The rationale for this subgroup definition is to have roughly 50 % of the subjects in each of the two subgroups of change from first aflibercept administration to AZURE baseline. In the VIEW2 study the median change from first injection to week 52 was 9 letters.

6 Statistical Methodology

6.1 Population Characteristics

6.1.1 Study Periods and Sample Sizes

Study sample sizes will be provided by trial unit and overall for all enrolled subjects.

The number of minor and major protocol deviations will be summarized by trial unit and overall for all enrolled subjects.

The number of major protocol deviations will be summarized by category and overall for all randomized subjects.

Screening Failures will only be listed together with the reason for their failure and all available disposition data (Date of informed consent, Reason for screen failure, Date of last visit).

For screening failures with an SAE all information related to the SAE will be listed.

6.1.2 Subject Validity Status

An overview-table for all randomized subjects will be given, displaying in each treatment group and overall the number and percentages of subjects:

- Subjects valid for safety analysis
- Subjects valid for FAS analysis
- Subjects valid for per protocol analysis
- Excluded from safety analysis
    - Never took study drug
- SAF, but excluded from FAS analysis
    - *[Reason, category from PDD]*
- FAS, but excluded from per protocol analysis
    - *[Reason, category from PDD]*

6.1.3 Subject Disposition

All enrolled subjects will be summarized together with the reason(s) for not entering the treatment period. Possible reasons are:

- Adverse event
- Death
- Withdrawal by subject
- Lost to follow-up

- Screen failure
- Other

An overview-table for all randomized subjects will be given, displaying in each treatment group and overall the number and percentages of subjects at end of treatment:

- Completed treatment period
- Not completed treatment period
    - withdrawn up to week 36
    - withdrawn after week 36 up to week 52
    - withdrawn after week 52 up to week 76
- Primary reason for not completing treatment
    - Adverse Event
    - Death
    - Withdrawal by subject
    - Lost to follow-up
    - Protocol Violation
    - Study terminated by sponsor
    - Treatment failure
    - Physician decision
    - Pregnancy
    - Other

A similar table will be provided displaying the result at end of study.

The subject disposition will be displayed for all enrolled subjects in a summary table showing the number of subjects randomized but not treated.

6.1.4 Demographics and Baseline Characteristics

Demographic variables will be summarized by treatment group and all treatment groups combined for all three analysis populations, depending on the type of data as described in Section 4.1.

The following demographic variables are recorded at screening. In case of repeated measures the last available value before randomization will be used for summary tables.

- sex (m/f/childbearing potential)
- race
- ethnicity
- age
- age (classified)
- weight
- height
- body mass index (BMI)

- study eye (left or right)
- cigarettes smoking (never/former/current)
- other tobacco smoking (never/former/current)

The treatment group comparability will be checked for each of the analysis populations mentioned above. This comparison will be done with respect to age by a one-way analysis of variance with treatment group as fixed factor and with respect to gender by a Chi-squared test.

Baseline characteristics will be summarized by treatment group and all treatment groups combined for all three analysis populations, depending on the type of data as described in Section 4.1.

The total score and sub-scores of the National Eye Institute 25-item Visual Function Questionnaire (NEI VFQ-25) will be calculated according to the NEI VFQ-25 scoring algorithm, August 2000 version. The most important instructions are displayed in appendix 8.1.

The following baseline characteristics will be summarized:

- Baseline BCVA letter scores (study eye)
- Baseline central retinal thickness (CRT) (study eye)
- Baseline choroidal neovascularization (CNV) area (study eye)
- Baseline CNV lesion type (study eye)
- Baseline total CNV lesion area (study eye)
- Baseline NEI VFQ-25 total score

The treatment group comparability will be checked for each of the analysis populations mentioned above. This comparison will be done with respect to

- Baseline BCVA letter score
- Baseline NEI VFQ-25 total score
- Baseline CRT
- Baseline CNV area

by a one-way analysis of variance with treatment group as fixed factor.

Results from prothrombin time (PT/INR) and partial thromboplastin time (PTT) measurements and physical examination at screening will only be listed.

The subgroups

- sex,
- geographic region and
- change of ETDRS BCVA letter score from first aflibercept administration to AZURE baseline

will be displayed for all three analysis populations by treatment group and overall.

6.1.5 Medical and Ophthalmological History

The number and percentages of subjects with the respective medical history will be displayed for SAF and FAS by treatment group and overall, stratified for ocular medical history in the study/fellow eye.

For the tables, subjects may be counted under multiple system organ classes (SOC) and preferred terms (PT), but for each SOC and PT, subjects are only counted once. Medical histories will be sorted by descending frequency, in all treatment groups combined, by the SOCs and within each SOC by PT.

The duration of nAMD (years) will be calculated per eye as described in Section 4.5. The variables recorded on the respective page of the CRF will be summarized by treatment group and all treatment groups combined for SAF and FAS, depending on the type of data as described in Section 4.1.

6.1.6 Prior and Concomitant Medications

Summaries of all prior and concomitant medications recorded will be presented in tabular form using 3-digit Anatomical Therapeutic Chemical (ATC) classification codes and preferred drug name via the World Health Organization Drug classification Dictionary (WHO-DD), latest version available before database lock, for SAF.

The medications will be classified as:

- Concomitant: Medications that are ongoing at, began after the start of study drug, or medications that were started after end of study drug.
- New Concomitant: Medications that began after the start of study drug, and those that were started after end of study drug.
- Prior: Medications that started and stopped before the start of study drug.

For each of these categories a table will be created, stratified by

- ocular medication (yes/no) and
- eye (study eye/fellow eye)

consisting of medication class and preferred name, sorted by descending frequencies. Bilateral medications are counted as medications for the study as well as the fellow eye.

Subjects may be counted under multiple medication classes and preferred names but within each category subjects are only counted once. However, multiple ATC codes per drug are

possible. Therefore, the same drug may be counted in more than one category for the same subject.

6.1.7 Exposure

The number of injections (counts and metrics) will be displayed for all three analysis populations by treatment group and overall.

Treatment duration is calculated as (date of 'end of treatment' - date of 'first injection after randomization' + 56)/7. The rationale for this is that the minimal time between aflibercept injections is 8 weeks that means 56 days. It will be summarized by treatment group and total for all three analysis populations.

The last actual injection interval will be calculated as (last injection – previous injection +1)/7 tabulated in both treatment groups and total for all three analysis populations – as counts, classified and metric variable as described in Section 4.1. The last intended injection interval as documented by the investigator will be summarized similarly.

6.2 Efficacy

Efficacy variables will be summarized by treatment group for each visit depending on the type of data as described in Section 4.1 for the FAS.

The analysis of the primary, key secondary, and secondary variables will be based on the results at Week 52. LOCF will be used for missing post-baseline values at 52 weeks[2].

Sensitivity analysis using RMM, OC and MI will be performed on the primary efficacy variable on the FAS. Analyses on the PPS and on subgroups will be considered as supplementary.

The efficacy variables and the ranking of their statistical analyses at the different timepoints (primary, [key] secondary, exploratory) are specified in Table 2. The first 9 efficacy variables and the ranking of their statistical analyses were specified in the study protocol. Additionally, "time to withdrawal" will be analysed in an exploratory manner.

Table 2: Efficacy variables

Item	Variable	Ranking	
		Week 52	Week 76
1	Change from baseline in ETDRS BCVA letter score for the study eye	Primary	Exploratory
2	Proportion of subjects maintaining vision (i.e. loss of < 15 letters) in the study eye	Key secondary	Exploratory
3	Proportion of subjects who gained from baseline ≥ 5 letters in the study eye	Secondary	Exploratory
4	Mean change from baseline in CRT in the study eye	Secondary	Exploratory
5	Mean change from baseline in CNV area in the study eye	Secondary	Exploratory
6	Proportion of subjects who lost ≥ 30 letters	Secondary	Exploratory
7	Mean change from baseline in total score for NEI VFQ-25	Secondary	Exploratory
8	Proportion of subjects for whom the treatment interval was extended	Exploratory	Exploratory
9	Total number of intravitreal injections required in the study eye	Exploratory	Exploratory
10	Time to withdrawal	Exploratory	Exploratory

BCVA = best corrected visual acuity; CNV: choroidal neovascularization;
CRT = central retinal thickness;
ETDRS = Early Treatment Diabetic Retinopathy Study;
NEI VFQ-25 = National Eye Institute 25-item Visual Function Questionnaire

[2] LOCF is considered an appropriate method for replacement of missing values in the primary and secondary efficacy analyses in the late phase of the study (i.e. carrying forward the last measurement taken between Week 36 to Week 48 if the Week-52 measurement is missing), even though nAMD is a progressive disease.

Subgroup analysis will be performed by sex, geographic region and change of ETDRS BCVA letter score from first aflibercept administration to AZURE baseline for

- Item 1: Change from baseline in ETDRS BCVA letter score for the study eye (LOCF),
- Item 2: Proportion of subjects maintaining vision (LOCF), and
- Item 7: Mean change from baseline in total score for NEI VFQ-25 (OC)

at Week 52 on FAS.

6.2.1 Primary Efficacy Variable Analysis

The primary efficacy variable is the change from baseline in ETDRS BCVA letter score for the study eye. The primary efficacy variable analysis will be performed on the FAS (defined in Section 5).

Statistical testing will be conducted to prove the non-inferiority of the extended-dosing regimen to the 2Q8 fixed dosing regimen.

The corresponding null hypothesis is H_0: $\mu_1 \leq \mu_2$-D versus the alternative hypothesis H_1: $\mu_1 > \mu_2$-D, where

D = non-inferiority margin

μ_i = mean change in BCVA letter score for the study eye from baseline to Week 52 in treatment group i.

i = 1: extended-dosing regimen
2: fixed-dosing regimen

The non-inferiority margin is 5 letters.

The methodological approach will be the calculation of a two-sided 95% confidence interval for the difference in the least squares (LS) means (extended-dosing group minus 2Q8 group) of the change in ETDRS letter score from study baseline to 52 weeks based on a one-way analysis of covariance (ANCOVA) with baseline measure as a covariate and treatment group as a fixed factor.

In the data analysis, ANCOVA can be performed using SAS PROC MIXED procedure with the following key SAS statements:

```
PROC MIXED DATA=<input dataset>;
     CLASS <treatment group>
     MODEL <change from baseline in BCVA>
           =<treatment group> <baseline BCVA>/DDFM=KenwardRoger;
     LSMEANS <treatment group>/
              PDIFF=CONTROLU('<Fixed dosing regimen>') CL
     ALPHA=0.025;
RUN;
```

The extended-dosing regimen will be considered to be non-inferior to the fixed 2Q8 regimen if the confidence interval of the difference lies entirely above -5 letters, where a positive difference favours extended dosing. A non-inferiority margin of 5 letters is consistent with margin used in the "CATT Study" (3).

If, and only if, the extended-dosing regimen is statistically proven to be non-inferior to the fixed 2Q8 regimen in the primary efficacy analysis, confirmatory testing will be continued to prove the non-inferiority of extended dosing to fixed 2Q8 with regard to the key secondary efficacy variable (maintenance of vision) in the FAS (with LOCF for missing 52-week ETDRS letter score).

As supplemental analysis the primary efficacy analyses will be repeated on the PPS (defined in Section 5) using LOCF. RMM, MI and OC in the FAS will serve as sensitivity analysis. Subgroup analysis are supplementary.

For illustrational purposes figures containing the mean ETDRS BCVA letter score changes from baseline together with their standard deviations over the course of time will be presented for LOCF in FAS.

6.2.2 Key Secondary Efficacy Variable

The key secondary efficacy variable is the proportion of subjects maintaining vision (i.e. loss of < 15 letters) in the study eye. The key secondary efficacy variable analysis will be performed on the FAS.

The corresponding null hypothesis is H_0: $p_1 \leq p_2-\Delta$ versus the alternative hypothesis H_1: $p_1 > p_2-\Delta$, where

p_i = proportion of subjects maintaining vision at Week 52 of treatment group i

Δ = pre-specified non-inferiority margin of 7%[3]

i = 1: extended-dosing regimen,
2: fixed-dosing regimen

The methodological approach will be the calculation of two-sided 95% confidence intervals using normal approximation of the difference between the proportions (extended-dosing group minus 2Q8 group) of subjects maintaining vision.

The extended-dosing regimen will be considered to be non-inferior to the fixed 2Q8 regimen if the confidence interval of the difference lies entirely above -7%, where a positive difference

[3] The non-inferiority margin of 7% was proposed by the EMA in their scientific advice of May 2007 during the discussion of the proposed 10% non-inferiority margin in the VIEW studies (EMEA/CHMP/SAWP/310870/ 2007 pages 22 and 23).

favours extended dosing. This conditional sequence of statistical hypotheses (a-priori ordered hypotheses) will control for multiplicity in the confirmatory analyses.

The analyses will be repeated on the PPS.

6.2.3 Other Secondary Efficacy Variables

The other secondary efficacy variables will be analysed in a descriptive manner. This may include 95% confidence intervals (CI) for treatment differences at Weeks 52 in an exploratory way. Details for calculation are given in Section 8.2.

The LOCF-analyses will be carried out on the FAS.

6.2.3.1 Proportion of Subjects who gained from BL ≥5 Letters in the Study Eye

BCVA will be measured at each visit up to the final visit. The changes from baseline will be calculated and summarized as described in Section 4.1, together with the respective 95%-CI (normal approximation) in each visit.

Subsequently, the BCVA results will be dichotomized in gain from baseline in the study eye of $\geq$ 5 letters or < 5 letters. The numbers and proportions of subjects with a gain $\geq$ 5 letters will be displayed at each visit together with the respective 95% CI (exact binomial) for both treatment groups and overall.

6.2.3.2 Mean Change from BL in CRT in the Study Eye

Optical coherence tomography (OCT) to measure the CRT will be performed at each visit up to the final visit.

The changes to baseline will be calculated and summarized together with the respective 95%-CI (normal approximation) at each visit.

For illustrational purposes figures containing the mean values of the changes from baseline together with their standard deviations over the course of time will be presented.

6.2.3.3 Mean Change from BL in CNV Area in the Study Eye

OCT to measure the CNV area will be performed at each visit up to the final visit. The changes to baseline will be calculated and summarized together with the respective 95%-CI (normal approximation) at each visit.

For illustrational purposes figures containing the mean values of the changes from baseline together with their standard deviations over the course of time will be presented.

6.2.3.4 Proportion of Subjects who lost ≥30 Letters in the Study Eye

The changes from baseline of the BCVA will be dichotomized in loss ≥ 30 letters in the study eye, and < 30 letters. The numbers and proportions of subjects with a loss ≥ 30 letters will be displayed at each visit together with the respective 95% CI (exact binomial) for both treatment groups.

6.2.3.5 Mean Change from BL in Total Score for NEI VFQ-25

NEI VFQ-25 will be evaluated at baseline, visit Week 52 and Week 76/final visit. LOCF will not be performed. From the 25 questions a total score and up to twelve sub-scores can be derived (see Section 8.1,Table 5). The (sub-) scores range from 0 to 100.

The changes from baseline of the total score will be summarized together with the respective 95%-CI (normal approximation).

For illustrational purposes figures containing the mean total values of the changes from baseline together with their standard deviations over the course of time will be presented.

6.2.4 Exploratory Efficacy Variables

The exploratory efficacy variables from will be analysed in a descriptive manner. This may include 95% confidence intervals for treatment differences at Weeks 52 and 76 in an exploratory way.

If not stated otherwise, the analyses will be carried out on the FAS and only observed cases (without LOCF) will be analysed.

6.2.4.1 Proportion of Subjects for whom the Treatment Interval was extended

A scheduled extended treatment interval (=extension) is defined as a scheduled interval of at least 9 weeks (+/- 3 days) between two injection visits, regardless of the real treatment interval.

An extension based on actual values is defined as an interval of at least 9 weeks between two injection visits.

A subject year will be calculated as the individual treatment duration divided by 365 days.

The following data will be tabulated for the extended-dosing group.

- Number of scheduled extensions at Weeks 52 and 76
- Number of scheduled extensions for study completers

The following data will be summarized on both treatment groups.

- Number of extensions based on actual values at Weeks 52 and 76
- Number of extensions based on actual values for study completers
- Number of extensions based on actual values, per subject year

6.2.4.2 Total number of ivt Injections required in the Study Eye

The total number of ivt injections required in the Study Eye is defined as the number of administered injections per subject.

Compliance (%) is calculated as the number of ivt injections in the study eye at the scheduled injection visits (+/- 3 days) divided by number of scheduled injection visits (multiplied by 100).

The following variables will be tabulated in both treatment groups:

- Number of injections at Weeks 52 and 76
- Number of observed injections at Weeks 52 and 76, per subject year
- Compliance at Weeks 52 and 76
- Number of injections per study completer

6.2.4.3 Week 76 Analysis

As described in Sections 0 to 6.2.3.5 the results at Week 52 are used to analyze the primary, key secondary or secondary variables.

If not specified otherwise, the analyses with respect to Week 76 will be performed in the same way as for Week 52 (except subgroup analysis), however in a descriptive manner.

6.3 Pharmacokinetics / Pharmacodynamics

Not applicable

6.4 Safety

The safety analysis will be conducted in the safety analysis set (as defined in Section 5).

6.4.1 Adverse Events (AEs)

The definitions of AEs and serious AEs (SAEs) are provided in the study protocol, Section 9.6.1.1. The classifications according to seriousness, intensity, causality, action taken, other specific treatment and outcome are provided in the study protocol, Section 9.6.1.2.

Treatment-emergent adverse events (TEAEs) are AEs that start after the first application of aflibercept in the study and less or equal than 30 days after the last dose of study drug.

Potential arterial thrombotic events (ATEs) will be evaluated by a masked adjudication committee according to criteria formerly applied and published by the Anti-Platelet Trialists' Collaboration (APTC) (2). Further details are described in the adjudication committee charter.

Overview tables by treatment and overall will be produced for the following categories for all AE and TEAE, presenting number of subjects and events:

- Subjects with at least one AE/TEAE
- Subjects with at least one ocular AE/TEAE
- Subjects with at least one ocular AE/TEAE in the study eye
- Subjects with at least one ocular AE/TEAE in the fellow eye
- Subjects with at least one non-ocular AE/TEAE
- Subjects with at least one (treatment emergent) APTC event
- Subjects with at least one AE/TEAE causally related to aflibercept
- Subjects with at least one AE/TEAE causally related to ivt injection
- Subjects with at least one AE/TEAE causally related to other protocol- required procedures
- Maximum intensity for any AE/TEAE
- Maximum intensity for any TEAE causally related to aflibercept
- AE/TEAE with outcome Death
- Subjects with at least one serious AE/TEAE
- Subjects with at least one serious TEAE causally related to aflibercept
- Subjects with at least one serious TEAE causally related to ivt injection procedures
- Subjects with at least one serious TEAE causally related to protocol- required procedures
- Discontinuation of study drug due to TEAE
- Discontinuation of study drug due to (treatment emergent) SAE
- Subjects with at least one non-treatment emergent AE (pre-treatment)

- Subjects with at least one non-treatment emergent AE (post-treatment)

The overview table for TEAEs will also be displayed stratified for the subgroups.

The following tables will present the respective AEs or TEAEs by MedDRA preferred term (PT) within primary system organ class (SOC) and summarized by treatment groups. Subjects may be counted under multiple system organ classes and preferred terms, but for each system organ class and preferred term, subjects are only counted once.

- TEAEs
- Ocular TEAEs
- Ocular TEAEs in the study eye
- Ocular TEAEs in the fellow eye
- Non-ocular TEAEs
- Serious TEAEs
- Serious ocular TEAEs in the study eye
- Serious ocular TEAEs in the fellow eye
- Serious non-ocular TEAEs
- TEAEs causally related to aflibercept
- TEAEs causally related to injection procedure
- TEAEs causally related to other procedures required by the protocol
- Serious TEAEs causally related to aflibercept
- TEAEs by maximum intensity
- Serious TEAEs by maximum intensity
- Serious TEAEs causally related to aflibercept by maximum intensity
- TEAEs by worst outcome
- Serious TEAEs by worst outcome
- Treatment emergent ATPC events[4]
- TEAEs resulting in discontinuation of aflibercept
- Treatment emergent Deaths
- Non-Serious TEAEs
- Non-treatment-emergent AEs
- Non-treatment-emergent AEs (pre-treatment)
- Non-treatment-emergent AEs (post-treatment)

The number and percentages of subjects affected as well as the number of events (except for those by maximum severity or worst outcome) will be displayed. SOCs will be sorted by descending frequency of all subjects, within each SOC the PT will be sorted by descending frequency of all subjects affected.

[4] APTCs are summarized by APTC-term and preferred term.

Additional the following subject listings will be provided.

- Serious adverse events
- Adverse events resulting in discontinuation of study drug
- Deaths

6.4.2 Laboratory Evaluations and Pregnancy Tests

All laboratory tests will be performed at a central laboratory at screening and final visit. The last measurement before first application of study-drug in this study will be used as baseline-value. Safety laboratory parameters to be evaluated are shown in Table 3.

Table 3: Laboratory safety parameters

Chemistry	**Urinalysis**	**Hematology**
Sodium	Glucose	Hemoglobin
Potassium	Protein	Hematocrit
Chloride	Specific Gravity	Red blood cell count
Calcium	Blood	Mean corpuscular volume (MCV)
Glucose	Ketones	Mean corpuscular hemoglobin concentration (MCHC)
Albumin	Protein: Creatinine Ratio (UPCR)	Mean corpuscular hemoglobin (MCH)
Total Protein, Serum		Leucocyte count
Creatinine		Differential count
Blood urea nitrogen (BUN)	**Coagulation**	Neutrophils
Aspartate aminotransferase (AST)	Prothrombin time (PT)	Lymphocytes
Alanine aminotransferase (ALT)	Partial thromboplastin time (PTT)	Monocytes
Alkaline phosphatase	International normalized ratio (INR)	Basophils
Total bilirubin		Eosinophils
Amylase		Platelet count
Serum pregnancy test for women of childbearing potential		

According to current International Conference on Harmonization guidelines, deviations from the reference range should be evaluated for clinical significance in each individual case. The reference ranges and the units and methods for all variables will be provided by the central laboratory.

Deviations of laboratory values from the laboratory reference ranges will be flagged on the laboratory listings.

Treatment-emergent laboratory abnormalities, defined as laboratory abnormalities after randomization under this study protocol, will be summarized by laboratory category and treatment group.

Summary statistics for the laboratory values will only be provided, when they are converted in comparable SI values. Values below "Normal Range Lower Limit" (LL) will be set to 0.5*LL for the calculation of summary statistics. Values above "Normal Range Higher Limit" (HL) will be set to HL.

Pregnancy tests will be performed for women of childbearing potential; a serum pregnancy test will be performed at screening and urine tests will be performed at baseline and Week 76. Beyond that schedule, pregnancy tests are to be done in women of childbearing potential as frequently as requested.

Results of the pregnancy tests for women of childbearing potential will be listed only.

6.4.3 Further Safety

6.4.3.1 Vital Signs

Vital signs (body temperature, blood pressure, pulse) including changes from baseline will be displayed for each treatment group.

6.4.3.2 Surgeries after Start of Study

Separate tables for treatment-emergent surgeries will be presented, overall and split by study eye and fellow eye.

6.4.3.3 Intra-Ocular Pressure (IOP)

At visits with study drug administration, pre-injection and 30 to 60-min post-injection IOP will be assessed for the study eye. All IOP measurements will be classified as follows:

- > 25 mmHg
- ≥ 30 mmHg
- ≥ 35 mmHg
- ≥ 40 mmHg

Any increase from pre-injection/baseline of ≥ 10 mmHg will be flagged.

The IOP analysis will include

- original measurements
- changes from pre-injection
- changes from baseline
- classified measurements

- flagged measurements

This analysis will be performed descriptively at all visits depending on the type of data as described in Section **Error! Reference source not found.**.

Additionally, the changes of the measurements between pre- and post-injection will be analyzed descriptively depending on the type of data as described in Section 4.1.

6.4.3.4 Slit Lamp Biomicroscopy

The slit lamp examination will be performed according to local medical practice and applicable medical standards at the site.

Frequency tables for normal/abnormal slit lamp biomicroscopy findings as well as for the grading of 'anterior chamber flare', 'anterior chamber cells' and 'anterior vitreous cells' will be provided.

6.4.3.5 Indirect Ophthalmoscopy

At visits with study drug administration, pre-injection indirect ophthalmoscopy will be assessed for the study eye. Post-injection assessments will be carried out 30-60 minutes following the IVT injection.

Frequency tables for normal/abnormal findings in the pre-injection assessment will be provided. The cup-to-disc ratio will presented by summary statistics. All other data will be listed.

Document History and Changes in the Planned Statistical Analysis

Main changes to „SAP final version 1.0, dated 29 May 2015“:

- In Section 4.5, the time-windows are corrected and pooling rules are added. Definition of geographic region
- In Section 6.1.4, diabetic status (yes/no), anti-hypertensive medication (yes/no) are deleted, because these parameter will not be collected, geographic region was added.
- In Section 6.1.6, subgroup analysis was added.

Main changes to „SAP final version 2.0, dated 08 September 2015“:

- Section 4.3.1: Adverse Events started later than 30 days after last dose of study-drug are not considered as treatment emergent.
- Section 4.3.2: RMM and OC were added. SAS-Code for MI-procedure was clarified.
- Section 4.3.3: An algorithm was given, to derive the flags for incomplete start- and stop-dates for concomitant medications.
- Section 4.5: Time-windows are slightly updated to be consistent with VIOLET.
- Section 5.1: Major protocol deviations have been adapted to version 8.0 of the PDD.
- Section 5.2: Definition of subgroups was added, including the additional subgroup “Change of the ETDRS BCVA letter score from first aflibercept administration to AZURE baseline (< 10 letters or >= 10 letters)”.
- Section 6.1: Clarifications and updates for analyses according to Bayer standards.
- Section 6.2.1: Clarifications and updates for analyses:
    - Updating SAS-Code for primary analysis
    - Adding sensitivity analyses RMM and OC
- Section 6.2.2: Clarifications and updates for analyses:
    - Dropping sensitivity analyses for key secondary parameter
- Section 6.2.3: Clarifications and updates for analyses:
    - Dropping sensitivity and PPS-analyses for other secondary parameter
    - Clarifying analyses for NEI-VFQ-25, especially dropping sub-score analysis
- Section 6.2.4: Clarifications and updates for analyses:
    - Clarifying analysis of treatment-extension

- Adding Compliance (%)

- Section 6.4: Clarifications and updates for analyses according to Bayer standards.
- Section 6.4.1: Adding overview TEAE tables stratified for subgroups.
- The analysis of the further safety measurements were described. The summaries of the fellow-eye was dropped.
- Formula for calculation of confidence intervals are given in the Appendix 8.1.

7 References

1. Aiello LP, Avery RL, Arrigg PG et al. Vascular endothelial growth factor in ocular fluid of patients with diabetic retinopathy and other retinal disorders. N Eng J Med 1994; 331:1480-1487.
2. Antiplatelet Trialists' Collaboration (2002) Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. BMJ, 324,71–86.
3. CATT Research Group. Ranibizumab and Bevacizumab for Neovascular Age-Related Macular Degeneration. N Eng J Med 2011; 365:1897-1908.
4. Mangione CM, MD NEI VFQ-25 Scoring Algorithm – August 2000, Version 2000, The National Eye Institute 25-Item, Visual Function Questionnaire (VFQ-25)
5. Rahmani B, Tielsch JM, Katz J et al. The cause-specific prevalence of visual impairment in an urban population. The Baltimore Eye Survey. Ophthalmology. 1996; 103:1721-1726.

8 Appendix

8.1 Calculation of the NEI VFQ-25 Scores

The calculation for NEI VFQ-25 sub-scale scores and total score will be performed according to the "NEI VFQ-25 Scoring Algorithm – August 2000 (4). The most important instructions are displayed below:

The NEI VFQ-25 consists of a base set of 25 vision-targeted questions representing 11 vision-related constructs, plus an additional single-item general health rating question, which are also presented in the protocol in Section 16.2.

The NEI VFQ-25 generates the following vision-targeted sub-scales (number of questions):

- global vision rating (1),
- difficulty with near vision activities (3),
- difficulty with distance vision activities (3),
- limitations in social functioning due to vision (2),
- role limitations due to vision (2),
- dependency on others due to vision (3),
- mental health symptoms due to vision (4),
- driving difficulties (3),
- limitations with peripheral (1),
- color vision (1), and
- ocular pain (2).

Additionally, the VFQ-25 contains the single general health rating question which has been shown to be a robust predictor of future health and mortality in population-based studies.

Scoring VFQ-25 is a two-step process:

1. First, original numeric values from the survey are re-coded following the scoring rules outlined in Table 4. All items are scored so that a high score represents better functioning. Each item is then converted to a 0 to 100 scale so that the lowest and highest possible scores are set at 0 and 100 points, respectively. In this format scores represent the achieved percentage of the total possible score, e.g. a score of 50 represents 50% of the highest possible score.
2. In step 2, items within each sub-scale are averaged together to create the 12 sub-scale scores. Table 5 indicates which items contribute to each specific sub-scale. Items that are left blank (missing data) are not taken into account when calculating the scale

scores. Sub-scales with at least one item answered can be used to generate a sub-scale score. Hence, scores represent the average for all items in the subscale that the respondent answered.

To calculate an overall composite score for the VFQ-25, simply average the vision-targeted subscale scores, excluding the general health rating question. By averaging the sub-scale scores rather than the individual items we have given equal weight to each sub-scale, whereas averaging the items would give more weight to scales with more items.

Table 4: Scoring Key: Recoding of Items

Item Numbers	Change original response category[a]	To recoded value of:
1,3,4,15c[b]	1 2 3 4 5	100 75 50 25 0
2	1 2 3 4 5 6	100 80 60 40 20 0
5,6,7,8,9,10,11, 12,13,14,16,16a	1 2 3 4 5 6	100 75 50 25 0 *
17,18,19,20,21, 22,23,24,25	1 2 3 4 5	0 25 50 75 100

(a) Pre-coded response choices as printed in the questionnaire.
(b) Item 15c has four-response levels, but is expanded to a five-levels using item 15b.
Note: If 15b=1, then 15c should be recoded to "0"
If 15b=2, then 15c should be recoded to missing.
If 15b=3, then 15c should be recoded to missing
* Response choice "6" indicates that the person does not perform the activity because of non-vision related problems. If this choice is selected, the item is coded as "missing."

Table 5: Averaging of Items to Generate VFQ-25 Sub-Scales

Scale	Number of items	Items to be averaged (after recoding per Table 3)
General Health	1	1
General Vision	1	2
Ocular Pain	2	4, 19
Near Activities	3	5, 6, 7
Distance Activities	3	8, 9, 14
Vision Specific: Social Functioning Mental Health Role Difficulties Dependency	 2 4 2 3	 11, 13 3, 21, 22, 25 17, 18 20, 23, 24
Driving	3	15c, 16, 16a
Color Vision	1	12
Peripheral Vision	1	10

Scoring example

Items 5, 6, and 7 are used to generate the near activities sub-scale score (Table 4). Each of the items has 6 response choices.

- Response choice 6 indicates that the respondent does not perform the activity because of reasons that are unrelated to vision. If a respondent selects this choice, the answer is treated as missing and an average of the remaining items is calculated.
- Response choice 5 indicates that an activity is so difficult that the participant no longer performs the activity. This extremely poor near vision response choice is recoded to "0" points before taking an average of all three items.
- To score all items in the same direction, Table 4 shows that responses 1 through 5 for items 5, 6, and 7 should be recoded to values of 100, 75, 50, 25, and 0 respectively.
- If the respondent is missing one of the items, the person's score will be equal to the average of the two non-missing items.

Formula: $\text{Mean} = \frac{\text{(Score for each item with a non-missing answer)}}{\text{Total number of items with non-missing answers}}$

Example:

With responses converted: = (25 + 100 + 25) / 3= 50

Note: 100 = Best, 0 = Worst possible score.

8.2 Calculation of Confidence Intervals for Treatment Differences

8.2.1 Rates and Proportions

The confidence interval using normal approximation is defined as

$$CI = p_{ext} - p_{fix} \pm z_{\alpha/2} * SQRT(p_{ext}(1-p_{ext})/n_{ext} + p_{fix}(1-p_{fix})/n_{fix})$$

where

SQRT is the square root function

p_{ext} = success rate in the extended dosing group

p_{fix} = success rate in the fixed dosing group

n_{ext} = number of subjects in the extended dosing group

n_{fix} = number of subjects in the fixed dosing group

z_{α} = α quantile of the standard normal distribution.

No continuity correction is used in this normal approximation.

This confidence interval can be calculated with SAS PROC FREQ using the table options 'RISKDIFF' and 'ALPHA=0.05'.

8.2.2 Means

The 95% confidence interval for the mean is defined as

$$CI = m_{ext} - m_{fix} \pm t_{1-\alpha/2,n} * SE$$

where under the assumption of equal variances:

m_{ext} = mean (difference to baseline) in the extended dosing group

m_{fix} = mean (difference to baseline) in the fixed dosing group

$t_{1-\alpha,n}$ = α quantile of the t-distribution.

n = $n_{ext} + n_{fix}$

n_{ext} = number of subjects in the extended dosing group

n_{fix} = number of subjects in the fixed dosing group

SE = pooled Standard Error

Under the assumption of unequal variances (Behrens-Fisher-Problem) Satterthwaite's formula for calculating the CI is available.

Both confidence intervals can be calculated with SAS PROC TTEST using the option 'ALPHA=0.05'. There the assumption of equal variances in both dsoing-groups will be automatically tested using an F-test. If the respective p-value is below 0.05 than Satterthwaite's formula will be used, otherwise the CI using the pooled Standard Error, described above more detailed detail.