Parexel International

OcuTerra Therapeutics, Inc.

OTT166-201

OTT166-201 A Phase 2 Randomized, Double-Masked, Vehicle-Controlled, Multicenter Study to Evaluate the Safety and Efficacy of OTT166 Ophthalmic Solution in the Treatment of Diabetic Retinopathy (DR)

Statistical Analysis Plan

Version: 3.0

Parexel Project Number: 268475

SPONSOR SIGNATURE PAGE

Approved by

Signed:_____ Date:_____

David J. Tanzer, MD
OcuTerra Therapeutics
Chief Medical Officer

Parexel Signature Page

Signatures below confirm that the Statistical Analysis Plan was developed in accordance with SOP-GDO-WW-019 and that it is approved for release.

This document has been approved and signed electronically on the final page by the following:

	Signatory
Author	Kathleen Russell Project Role: Biostatistics Lead

TABLE OF CONTENTS

REVISION HISTORY

Version No.	Effective Date	Summary of Change(s)
1.0	12-May-2022	New document
2.0	30-June-2023	All the changes made in this version are described in Appendix 6.3
3.0	TBD	All the changes made in this version are described in Appendix 6.3

LIST OF ABBREVIATIONS

Abbreviation/Acronym	Definition/Expansion
AE	Adverse Event
AESI	Adverse Event of Special Interest
ANCOVA	Analysis of Covariance
Anti-VEGF	Anti-Vascular Endothelial Growth Factor
ASNV	Anterior Segment Neovascularization
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Curve
BCVA	Best Corrected Visual Acuity
BID	Twice Daily
BL	Baseline
BMI	Body Mass Index
CFP	Color fundus photography
CI	Confidence interval
CI-DME	Center-Involved Diabetic Macular Edema
CNV	Choroidal Neovascularization
CSR	Clinical Study Report
CST	Central Subfield Thickness
DME	Diabetic Macular Edema
DR	Diabetic Retinopathy
DRIL	Disorganization of Retinal Inner Layers
DRSS	Diabetic Retinopathy Severity Scale
eCRF	Electronic Case Report Form
EOS	End of Study
EOT	End of Treatment
ETDRS	Early Treatment Diabetic Retinopathy Study
ETRDR	Early Treatment Report Diabetic Retinopathy
FA	Fluorescein Angiography
FAS	Full analysis set
FAZ	Foveal Avascular Zone

Abbreviation/Acronym	Definition/Expansion
HbA1c	Glycosylated Hemoglobin
ICH	International Conference on Harmonisation
ICE	Intercurrent Event
IOP	Intraocular Pressure
IRMA	Intraretinal microvascular abnormality
ITT	Intent-to-Treat
IVT	Intravitreal Injection
MAR	Missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mantel-Haenszel
MV	Macular Volume
nAMD	Neovascular Age-Related Macular Degeneration
NPDR	Non-Proliferative Diabetic Retinopathy
OCT	Optical Coherence Tomography
OCTA	Optical Coherence Tomography – Angiography
OU	Both eyes
PDR	Proliferative Diabetic Retinopathy
PPS	Per-Protocol Set
PRP	Pan-retinal Photocoagulation
PT	Preferred Term
QID	Four Times Daily
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard deviation
SD-OCT	Spectral-Domain Optical Coherence Tomography
SE	Standard error
SE	Study eye
SI	Standard International
SoA	Schedule of Activities
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event

Abbreviation/Acronym	Definition/Expansion
TESAE	Treatment Emergent Serious Adverse Event
VEGF	Vascular Endothelial Growth Factor
VTC	Visually Threatening Complications

1 INTRODUCTION

OTT166 is a representative from a new class of drugs, i.e., a topical integrin inhibitor targeted to the back of the eye. It is being developed for treatment of Diabetic Retinopathy (DR). This study is designed to provide the proof of concept that non-proliferative diabetic retinopathy (NPDR) and mild proliferative diabetic retinopathy (PDR) can be effectively treated with a topical eye drop formulation of OTT166. Furthermore, this study is designed to determine the dose-exposure response of OTT166 in terms of efficacy and safety in order to support the dose selection for the subsequent pivotal Phase 3 studies with OTT166 in DR.

This Statistical Analysis Plan (SAP) describes all planned analyses for the Clinical Study Report (CSR) of Phase 2 Randomized, Double-Masked, Vehicle-Controlled, Multicenter Study to Evaluate the Safety and Efficacy of OTT166 Ophthalmic Solution in the Treatment of DR.

The structure and content of the SAP are based upon the International Conference on Harmonisation (ICH) E3 – Guideline for Industry Structure and Content of Clinical Study Reports.

The content of this SAP is based on following study document:

- Study Protocol Version 5.0 dated October 10, 2023.

2 STUDY OBJECTIVES AND ENDPOINTS

Primary Objectives:	**Primary Endpoints:**
To characterize the safety of topical OTT166 in participants with DR	• Proportion of participants who develop treatment-emergent adverse events (TEAEs) through Week 24
To characterize the efficacy of topical OTT166 in participants with DR	• Proportion of participants who have improved by ≥ 2 steps from baseline at Week 24 as determined by central reading center assessment using the DRSS
Secondary Objectives:	**Secondary Endpoints:**
To determine if topical OTT166 will prevent or delay the occurrence of visually threatening complications (VTC), defined as PDR worse than mild and/or anterior segment neovascularization (ASNV) determined by the investigator and/or CI-DME (defined below)	• Proportion of participants developing worse than mild PDR (DRSS 65 and above) at Week 24 • Time to development of PDR worse than mild (DRSS 65 and above) Proportion of participants who develop ASNV determined by the investigator at Week 24 • Proportion of participants who develop VTC at Week 24 • Time to development of PDR worse than mild (DRSS 65 and above) or CI-DME

To determine if topical OTT166 will prevent or delay the occurrence of CI-DME (CI-DME is defined as the presence of fluid in the central subfield in participants who have no fluid at baseline or CST> 325 µm)	• Proportion of participants who develop CI-DME at Week 24 • Time to development of CI-DME
To determine the effect of OTT166 on DRSS in participants with moderately severe to severe NPDR and mild PDR treated with topical OTT166	• Proportion of participants with change in DRSS steps at Week 24 compared to baseline – Change in DRSS steps is defined as DR worsening or improving by 1, 2, or ≥ 3 steps. • Proportion of participants with mild PDR (DRSS score 61B) at baseline who regress to NPDR (DRSS score ≤ 53) by Week 24
To determine the effect of OTT166 on BCVA in participants with moderately severe to severe NPDR and mild PDR	• Mean and median change in BCVA (ETDRS letters) from baseline to Week 24 • Lines gained/lost in BCVA (± 5, 10, and 15 ETDRS letters) at Week 24 • Mean and median change in area under the curve (AUC) for change in BCVA from baseline to Week 24
To determine the effect of OTT166 on central subfield thickness in participants with moderately severe to severe NPDR and mild PDR	• Mean and median change in CST from baseline to Week 24 • AUC for change in CST from baseline to 24 weeks
To determine the effect of OTT166 on the need for rescue therapy in participants with moderately severe to severe NPDR and mild PDR	• Proportion of participants who met the objective rescue criteria as defined in protocol Section 2.14.9. • Time to meet objective rescue therapy criteria as defined in protocol Section 2.14.9.
Exploratory Objectives	**Exploratory Endpoints**
To evaluate the change in DME in participants treated with OTT166	• Change in Macular Volume (MV) from baseline to Week 24 for 1-, 3-, and 6-mm regions

To assess the effects of treatment on anatomic markers other than DRSS and on retinal non-perfusion	• Change in area of retinal non-perfusion from baseline to Week 24 seen on widefield fluorescein angiography (FA) • Proportion of participants with intact foveal avascular zone (FAZ) seen on widefield FA at Week 24 • Optical coherence tomography angiography (OCTA) endpoints within the 6 mm area: - Change from baseline to Week 24 in FAZ area - Change from baseline to Week 24 in non-perfusion area - Change from baseline to Week 24 in vessel caliber - Ischemic Index at Week 24 compared to baseline - Proportion of participants with new neovascularization at Week 24 - Proportion of participants with new intraretinal microvascular abnormality (IRMA) at Week 24 - Change in vessel density from baseline to Week 24 in the superficial and deep layers

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

This is a Phase 2 randomized, double-masked, vehicle-controlled, multicenter study to evaluate the safety and efficacy of OTT166 ophthalmic solution in the treatment of DR. The OTT166 201 Phase 2 will be conducted to select an optimum dosing regimen (frequency) for Phase 3 pivotal trials.

The study population will be participants with Type 1 or Type 2 diabetes under adequate control as evidenced by glycosylated hemoglobin (HbA1c$\leqslant$12.0%), who show typical fundus alterations and have been diagnosed with moderately severe to severe NPDR or mild PDR, as defined by the DRSS Steps 47, 53 and 61B (Appendix 6.2). Participants may have fluid in the macula with BCVA of $\geqslant$69 Early Treatment Diabetic Retinopathy Study (ETDRS) letters (Snellen equivalent 20/40) and Central Subfield Thickness (CST) of $\leqslant$ 325 µm, and a normal foveal contour, which needs to be confirmed by an independent central reading center, in at least 1 eye.

Approximately 210 participants diagnosed with moderately severe to severe NPDR or mild PDR and who are treatment naïve (*i.e.*, no prior anti-Vascular Endothelial Growth Factor [VEGF] or laser (focal, grid, Pan-Retinal Photocoagulation [PRP]) administered) will be randomized 2:2:1:1 into the following groups: OTT166 5% BID, OTT166 5% QID, vehicle control 5% BID, vehicle control 5% QID. Randomization will be stratified by Screening DRSS score (47 or 53 or 61B). Participants with PDR (DRSS score 61B) will be capped at 20% of all randomized participants. 5% OTT166

ophthalmic solution or vehicle control, will be given to the participants who will self-administer one 50-microliter eye drop of study solution at the prescribed frequency based on randomization, for 24 weeks.

Eyedropper bottles containing active treatment of OTT166, and corresponding vehicle control are identical in appearance (size, color, shape) and associated with a unique number. Eye dropper bottles will be dispensed at the study visits summarized in the schedule of activities (SoA) in Appendix 6.1.

3.2 Endpoints

3.2.1 Primary Endpoint - Safety

The primary safety outcome measure of the study will be the proportion of participants who develop TEAEs through Week 24.

3.2.2 Primary Efficacy Endpoint and Estimand

The primary efficacy outcome measure of the study will be the proportion of participants who have improved by ⩾ 2 steps from baseline at Week 24 as determined by central reading center assessment using the DRSS. The primary clinical question of interest is to determine the effectiveness of different doses of OCT166 versus vehicle control on ⩾2 steps improvement from BL in DRSS score at Week 24 without requiring rescue treatment if the study drug is not discontinued prior to Week 24. This will support the dose selection for subsequent pivotal studies of OTT166. The estimand for the primary efficacy analysis is defined as follows:

	Primary Estimand
Treatment conditions of interest	OTT166 5 mg/day, OTT166 10 mg/day and matching vehicle control
Participant Population	DR participants with Type 1 or Type 2 diabetes under adequate control (HbA1c ⩽12.0%) who show typical fundus alterations and have been diagnosed with moderately severe to severe NPDR or mild PDR as defined by the DRSS Steps 47, 53, and 61B, with BCVA ⩾ 69 ETDRS letters, CST of ⩽ 325 µm, and a normal foveal contour.
Endpoint	≥2 steps improvement from BL in DRSS score at Week 24
Population level summary	Difference in proportion of participants achieving ≥2 steps improvement from BL in DRSS score at Week 24 between each OTT166 treatment group and the combined vehicle control group.

	Primary Estimand
Intercurrent events (ICEs) and strategies to handle ICEs	a) Rescue treatment in the study eye due to worsening DRSS level to 65 or higher prior to assessment of DRSS at Week 24 *Composite strategy (non-response)* b) Death prior to DRSS assessment at Week 24 *Hypothetical strategy* c) Use of rescue treatment in the study eye for any other reason prior to DRSS assessment at Week 24 *Composite strategy (non-response)* d) Premature discontinuation of study drug for any reason prior to DRSS assessment at Week 24 *Hypothetical strategy* e) Study drug compliance <75% through Week 24 *Treatment policy strategy* f) Study drug compliance between 75% to <90% through Week 24 *Treatment policy strategy*
	ICE(s) with a composite variable strategy take priority over ICEs with other strategies. And ICEs with a hypothetical strategy take priority over ICEs with treatment policy strategy.

A Supplemental Estimand will be defined and analyzed where the ICEs defined above will be handled using alternative strategies as follows: c and e are changed to Hypothetical strategy.

3.2.3 Secondary Endpoints

The secondary outcome measures are as follows:

- Proportion of participants developing worse than mild PDR (DRSS 65 and above) at Week 24.
- Proportion of participants who develop ASNV determined by the investigator at Week 24.
- Time to development of PDR worse than mild (DRSS 65 and above).
- Proportion of participants who develop CI-DME at Week 24.
    - CI-DME is defined as the presence of fluid in the central subfield in participants who have no fluid at baseline or CST> 325 µm.
- Time to development of CI-DME.
- Proportion of participants with change in DRSS steps at Week 24 compared to baseline.
    - Change in DRSS steps is defined as DR worsening or improving by 1, 2, or ≥3 steps.

- Proportion of participants with mild PDR (DRSS score 61B) at baseline who regress to NPDR (DRSS score ⩽ 53) by Week 24.
- Mean and median change in BCVA (ETDRS letters) from baseline to Week 24.
- Lines gained/lost in BCVA (± 5, 10, and 15 ETDRS letters) at Week 24.
- Mean and median change in Area Under the Curve (AUC) for change in BCVA from baseline to Week 24.
    - AUC from baseline to 24 weeks will be calculated by the linear trapezoidal method:

$$AUC_{(t_i - t_{i-1})} = (t_i - t_{i-1}) * \frac{f(t_i) + f(t_{i-1})}{2}$$

$$AUC_{(24-0)} = \sum_{t_0=0}^{24} AUC_{(t_i - t_{i-1})}$$

      where, t_i = # of weeks since baseline, which may include 0, 4, 8, 12, 16, 20, 24 according to schedule of visits, if available.
- Mean and median change in CST from baseline to Week 24.
- AUC for change in CST from baseline to Week 24.
- Proportion of participants who met the objective rescue criteria as defined in protocol section 2.14.9.
- Time to meet the objective rescue therapy criteria as defined in protocol section 2.14.9.
- Proportion of participants who develop VTC at Week 24
- Time to development of PDR worse than mild (DRSS 65 and above) or CI-DME

3.2.4 Exploratory Endpoints

Exploratory imaging outcome measures are as follows:

- Change in Macular Volume (MV) from baseline to Week 24 for 1, 3, and 6 mm regions.
    - MV for the 1 mm region is the central 1 subfield, for the 3 mm region is the middle 5 subfields and for the 6 mm region is all 9 subfields as provided by the imaging center
- Change in area of retinal non-perfusion from baseline to Week 24 seen on widefield Fluorescein Angiography (FA).
- Proportion of participants with intact foveal avascular zone (FAZ) seen on widefield FA at Week 24.
- Optical coherence tomography angiography (OCTA) endpoints within the 6mm area:
    - Change from baseline to Week 24 in FAZ area
    - Change from baseline to Week 24 in non-perfusion area
    - Change from baseline to Week 24 in vessel caliber
    - Change from baseline in Ischemic Index at Week 24 defined as total area of non-perfusion divided by total area of retina visualized (28.3 mm^2)

- Proportion of participants with new neovascularization at Week 24
- Proportion of participants with new IRMA at Week 24
- Change in vessel density from baseline to Week 24 in the superficial and deep layers

3.2.5 Pharmacokinetic, Pharmacodynamic and Immunogenicity Endpoints

In this study, no pharmacokinetic, pharmacodynamic or immunogenicity data are collected.

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures, and data listings to be included in the report will be independently checked for consistency, integrity, and in accordance with standard Parexel procedures.

4.2 General Presentation Considerations

'Baseline' is defined as the last available pre-treatment assessment for safety assessments. For DRSS score, the Baseline value is defined as the Screening/Visit 1 result. Pre-treatment values for DRSS were collected at the Screening/Visit 1 using fluorescein angiography in addition to fundus photography whereas Baseline/Visit 2 used fundus photography only. Participants were enrolled into the study and randomized to the study treatment based on the Screening/Visit 1 DRSS result. All other efficacy analyses that do not use DRSS score will use the last available pre-treatment assessment as Baseline. If an assessment is taken on the same day as first administration of study drug, it is eligible to be used as the Baseline value. This is considered acceptable as this measurement is still the best representation of the Baseline value of the given assessment since it is highly unlikely that the study drug could have an impact on any measurement in a short period of time.

'End of Study' is defined as the last available post-treatment assessment.

'Study Day' will be calculated relative to the date of randomization for untreated randomized participants and relative to the date of first dose for treated participants. If event is prior to the first dose, then study day is calculated as:

Date of Event – Date of Day 1.

If event is after the first dose, then study day is calculated as:

Date of Event – Date of Day 1 + 1.

'Study day 1' is defined as the date of randomization for untreated randomized participants or the date of first dose for treated participants.

Time to event analyses will be calculated from Study Day 1, including time to DRSS-related endpoints comparing to the Baseline value. These analyses will consider only post-Baseline assessments. The Visit 2/Baseline visit will not be considered.

Duration of Treatment (days) will be calculated as:

Last Date of Dose – First Date of Dose + 1.

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, upper quartile, lower quartile, minimum, maximum and number of observations, unless otherwise stated. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, median, lower quartile and upper quartile will be reported to one more decimal place than the raw data recorded in the database. The SD (and SE, if applicable)

will be reported to two more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.

Categorical data will be summarized in terms of the number of participants providing data at the relevant time point (n), frequency counts and percentages.

Percentages will be presented to one decimal place except 100%. 100% will be presented as 100% instead of 100.0%. Percentages will not be presented for zero counts. Percentages will be calculated using the number of participants (n) as the denominator. If sample sizes are small, the data displays will show the percentages, but any textual report will describe frequencies only. A missing category shall be included only for categorical variables with missing data. The missing category will be omitted if there were no missing values for that variable.

Changes from baseline in categorical data will be summarized using shift tables where appropriate.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. P-values less than 0.001 will be presented as "<0.001". P-values greater than 0.999 will be presented as ">0.999".

Confidence intervals (CIs) will be presented to one more decimal place than the raw data.

4.2.1 Handling of Incomplete Dates

Every effort will be undertaken to avoid any missing or incomplete data. If variables are imputed, the analysis dataset will contain a new variable with imputed value and the original variable will contain the original missing value.

Inevitable partial dates will be imputed for start/end dates of AEs or concomitant medications.

Partial AE Start Date

- If only day is missing, then it will be set to:
    - First day of the month that the AE occurred, if month/year of the AE start date is different than the month/year of the first dose date.
    - The day of first dose date, if month/year of the AE start date is the same as month/year of the first dose date and month/year of the AE end date is different.
    - The day of first dose date or day of AE end date, whichever is earliest, if the month/year of the AE start and month/year of the first dose date and month/year of the AE end date are the same.
- If only month is missing, then it will be set to the earliest of the following:
    - January, as long as this date is after the first dose date
    - Month of the first dose date, if this date is the same day and year that the AE occurred
- If the day and month are both missing, then it will be set to the earliest of the following:
    - January 1 of the year, as long as this date is after the first dose date
    - Month and day of the first dose date, if this date is the same year that the AE occurred
    - The AE end date
- If dates are completely missing, then no imputation will be made.

Partial AE End Date

- If only day is missing, then it will be set to the earliest of the last follow-up date, last day of the month, or the date of death.
- If only month is missing, then it will be set to the earliest of the last follow-up date, December, or the date of death.
- If year is missing or AE is ongoing, the end date will not be imputed.
- If dates are completely missing, then no imputation will be made.

If the imputed AE end date is before the corresponding AE start date, then the AE end date will be set to the AE start date.

The same rule above will be applied to the start and end dates of concomitant medications.

4.3 Software

All report outputs will be produced using SAS® version 9.4 or later in a secure and validated environment.

4.4 Study Participants

4.4.1 Analysis Sets

For purposes of analysis, analysis sets are defined as follows:

Analysis Set	Description
Enrolled Set (ES)	All participants who signed informed consent.
Intent-to-Treat Analysis Set (ITT)	All randomized participants who received at least one dose of the study drug. Participants will be included into the analysis as randomized. This corresponds to what is referred to in ICH E9 as the Full Analysis Set.
Per-Protocol Set (PPS)	All ITT participants without any major protocol deviations that could impact the primary endpoint.
Safety Analysis Set (SS)	All participants who receive at least one dose of study drug. Participants will be included into the analysis based on treatment actually received, regardless of the treatment randomized.

The ITT Analysis Set will be used for all efficacy analyses and BL data analyses except participant disposition which will be shown for the ITT Analysis Set and for the Safety Analysis Set if the two analysis sets differ. All randomized participants are expected to be dosed and any exceptions will be described in the CSR. The PP set will be determined prior to database lock and will be used as a supplementary analysis of the primary endpoint.

For analyses based on the SS, participants randomized to a control treatment group who receive any dose of active OTT166 treatment will be analyzed in the corresponding OTT166 treatment group (BID or QID).

A Data Review Meeting will be arranged to identify analysis sets or data to be included/excluded from analysis, including a review/confirmation of participants with ICEs. Decisions made regarding

the exclusion of participants and/or participants data from analyses will be made prior to unmasking and will be documented and approved by OcuTerra Therapeutics, Inc.

All safety analyses will be based on the Safety Analysis Set (SS).

If a participant is stratified incorrectly, the participant will be included into the analysis as stratified per the ITT principle.

The number and percentage of participants in each analysis set will be summarized by treatment group for the ES. Inclusion in the PP set will be indicated in the listing of participant disposition for the ITT Analysis Set.

Unless otherwise specified, participant disposition, demographics, and baseline characteristics will be summarized by the randomized treatment group (i.e., vehicle control doses (5% BID, 5% QID), vehicle control combined, active OTT166 doses (5% BID, 5% QID)), and overall.

4.4.2 Disposition of Participants

A clear accounting of the disposition of all participants who enter the study will be provided, from screening to study completion.

Participant disposition will be summarized for the ITT Analysis Set by treatment and overall, for:

- Number of participants randomized
- Number and percentage of participants received at least one dose of study drug
- Number and percentage of participants completed study drug
- Number and percentage of participants discontinued from study drug and primary reason for premature discontinuation
- Number and percentage of participants completed study
- Number and percentage of participants discontinued from study and primary reason for premature discontinuation

Participant disposition for participants who were screen failures (defined as participants in the Enrolled Set who were not randomized) will be summarized and will include the reason for screen failure.

By-participant listings of participant disposition details will also be provided.

4.4.3 Protocol Deviations

Protocol deviations will be classified as “major” or “minor” on ongoing basis by the clinical study team and sponsor.

Major protocol deviations are defined as those deviations from the protocol likely to have an impact on the perceived efficacy and/or safety of study treatments. The impact of major protocol deviations on the efficacy results will be investigated by assessing the robustness of the study results and conclusions to the choice of analysis set, both including and excluding data potentially affected by major protocol deviations.

Major protocol deviations and any action to be taken regarding the exclusion of participants or affected data from specific analyses are defined in the project-specific Protocol Deviation Specification. Final decisions made regarding the exclusion of participants and/or participants data

from analyses will be made prior to unmasking and will be documented and approved by OcuTerra Therapeutics, Inc.

The number and percentage of participants with at least one major protocol deviation as well as major protocol deviation categories for the ITT Analysis Set by treatment and overall will be summarized. This summary will also include a summary of participants excluded from the PPS due to major protocol deviations by category.

A by-participant listing of all protocol deviations will also be provided.

4.5 Demographics and Baseline Characteristics

The following demographics and baseline characteristics will be summarized on ITT Analysis Set by treatment and overall:

- Age (years)
- Gender
- Race
- Ethnicity
- Height (cm) at Baseline
- Weight (kg) at Baseline
- Body Mass Index (kg/m^2) at Baseline
- HbA1C at Baseline
- Intraocular pressure at Baseline
- DRSS steps at the Baseline Visit
- ETDRS letter score at Baseline
- Central Subfield Thickness (CST) at Baseline
- Duration of Diabetes (years) defined as time from diagnosis to randomization
- Diabetes Type (Type 1 and Type 2)
- Diagnosis of systemic hypertension

Additional demographic or baseline assessments may be added to the summary tables.

A separate summary of subgroups on ITT Analysis Set will also be provided by treatment and overall:

- Age (<65 years, ≥65 years)
- BMI (Normal [<25 kg/m^2], Overweight [≥25 to <30 kg/m^2], Obese [≥30 kg/m^2])
- HbA1C (≤7%, >7%)
- Duration of Diabetes (<median, ≥median)
- Baseline DRSS (47 or 53 or 61B)
- Presence of fluid in the central subfield at Baseline (Y/N)

By-participant listing of demographic and other baseline characteristics as well as subgroups will also be provided.

4.6 Medical and Surgical History

Medical and surgical history at screening will be collected under the Medical History form of the eCRF and will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 24.1 or higher. Medical and surgical history will be summarized for the Safety Analysis Set by treatment and overall using system organ class (SOC) and preferred term (PT).

A by-participant listing for medical and surgical history will be provided.

4.7 Prior and Concomitant Medications

Prior medication is defined as any medications that started and stopped before the first dose of study drug.

Concomitant medication is defined as any medications that started before the first dose of study drug and stops on or after the first dose of study drug, as well as any medications that started on or after the first dose of study drug.

Medication start and stop dates will be compared to the date of first dose of study drug to allow medications to be classified as either prior or concomitant. Medications starting after the End of Study, as previously defined, will not be classified or summarized.

Incomplete start and end date of the medications will be imputed as described in Section 4.2.1.

Medications will be coded using the World Health Organization Drug (WHO-Drug) Global March 2022 B3G. Prior and concomitant medications will be summarized on Safety Analysis Set by treatment and overall. Participants with more than one medication in a given Anatomical Therapeutic Chemical (ATC) level 4 and preferred term will be counted only once.

By-participant listings of both prior and concomitant medications will also be provided.

4.8 Rescue Therapy

Participants meeting one or all the following criteria in the study eye would qualify for rescue with anti-VEGF therapy and/or PRP laser after consultation with the Medical Monitor:

Objective Rescue Therapy Criteria:

- Worsening DRSS level to 65 or higher as determined by the central reading center
- Decrease in BCVA of $\geq$ 10 ETDRS letters from BL associated with presence of new fluid or an increase in fluid in the central subfield compared to baseline at any visit
- Decrease in BCVA of 5 to 9 ETDRS letters from BL at two consecutive visits 4 weeks apart associated with presence of new fluid or an increase in fluid in the central subfield compared to baseline
- Increase in CST of $\geq$ 75 microns from BL associated with presence of new fluid or an increase in fluid in the central subfield compared to baseline
- Development of anterior segment neovascularization (ASNV)

Subjective Rescue Therapy Criteria:

- Any other condition that, in the clinical opinion of the investigator, warrants rescue therapy after consultation with the Medical Monitor

Summary tabulations of number and percentage of participants with rescue therapy, reasons for rescue and type of rescue therapy will be provided.

A by-participant listing will be provided.

The start date of rescue therapy in the study eye will be determined by the earliest concomitant medication in the study eye with ATC4CD="SL01A" or the earliest ocular concomitant procedure confirmed by the clinical team to be rescue therapy. This will be subset on the participants who have completed the "Rescue Therapy for Ocular Diabetic Complications" eCRF.

4.9 Treatment Exposure and Compliance

4.9.1 Treatment Exposure

Exposure to study drug will be summarized using the ITT Analysis Set and also the Safety Analysis Set. The total number of doses administered, and the duration of treatment will be summarized in a summary table by treatment group and overall.

A by-participant listing of treatment exposure will be provided.

4.9.2 Treatment Compliance

Overall treatment compliance (%) will be calculated as:

$$\frac{\text{Total Number of Doses of Study Drug Administered}}{\text{Total Number of Doses of Study Drug Planned}} \times 100\%$$

The total number of doses of study drug administered is the actual number of drops administered. In the event a participant takes study drug after their EOS date, the total number of doses administered will not include any doses recorded after the EOS date.

The total number of doses of study drug planned is calculated based on a participant's duration in the study and the randomized treatment. A participant randomized to BID treatment will have 2 times the (date of EOS – date of first dose + 1). A participant randomized to QID treatment will have 4 times the (date of EOS – date of first dose + 1).

For participants who take rescue therapy in the study eye during the study, the total number of doses of study drug planned is calculated based on a participant's treatment duration prior to rescue and randomized treatment. A participant randomized to BID treatment will have 2 times the (date of EOT – date of first dose + 1). A participant randomized to QID treatment will have 4 times the (date of EOT – date of first dose + 1).

A summary of the treatment compliance will be provided for the ITT Analysis Set and also the Safety Analysis Set by treatment and overall.

Monthly (every 4 weeks) treatment compliance (%) will be calculated as:

$$\frac{\text{Total Number of Doses of Study Drug Administered each Month}}{\text{Total Number of Doses of Study Drug Planned for each month}} \times 100\%$$

The total number of doses of study drug planned is calculated based on a participant's duration in the study and the randomized treatment. A participant randomized to BID treatment will have 2 times the minimum of 28 or the (date of EOS – date of first dose in month + 1). A participant randomized to QID treatment will have 4 times the minimum of 28 or the (date of EOS – date of first dose in month + 1).

For participants who take rescue therapy in the study eye during the study, the total number of doses of study drug planned is calculated based on a participant's treatment duration prior to rescue and randomized treatment. A participant randomized to BID treatment will have 2 times the minimum of 28 or the (date of EOT – date of first dose in month + 1). A participant randomized to QID treatment will have 4 times the minimum of 28 or the (date of EOT – date of first dose in month + 1).

Number and percentage of participants with compliance in categories (<75%, 75-<90%, ≥90%) will be provided for overall and monthly treatment compliance.

A by-participant listing of treatment compliance will be provided.

4.10 Efficacy Evaluation

4.10.1 Analysis and Data Conventions

This study is designed to evaluate the superiority of two doses of OTT166 (5 mg/day and 10 mg/day) compared to vehicle control for the treatment of DR. For each dose within each endpoint, the null hypothesis is that there is no difference between OTT166 and vehicle control or the difference is in the direction of inferiority of OTT166, and the alternative hypothesis is that there is a difference between OTT166 and vehicle control in the direction of superiority of OTT166.

All statistical tests will be one-sided and will be performed at the 5% level of significance, unless otherwise stated. The vehicle control group will be combined for all efficacy analyses and two-sided 90% confidence intervals will be presented. A comparison between the two control groups will be performed for the primary endpoint to confirm the groups are comparable.

4.10.1.1 Adjustments for Covariates

The primary efficacy analysis will be adjusted for the randomization stratification factor, the Screening DRSS score (47 or 53 or 61B).

4.10.1.2 Handling of Missing Data and Data After Intercurrent Events

Primary Efficacy Analysis:

Handling of missing DRSS response at Week 24 data in participants without ICEs:

- Multiple imputation assuming missing at random (MAR).

Handling of DRSS response at Week 24 data missing after ICEs:

- Single imputation (non-response) when applying composite variable strategy.
- Multiple imputation assuming MAR when applying hypothetical and treatment policy strategy.

Handling of DRSS data response at Week 24 observed after ICEs:

- Replaced by non-response (single imputation) in the analysis when applying composite variable strategy.
- Replaced by MAR multiple imputation algorithm based on data from participants not affected by the respective ICE when applying hypothetical strategy.
- Kept in the analysis when applying treatment policy strategy.

If a participant experiences more than one ICE, ICE(s) with a composite variable strategy take priority over ICEs with hypothetical strategies, which take priority over ICEs with a treatment policy strategy.

MAR imputation will be generated based on a regression model for DRSS levels with covariates for randomized treatment group and the randomization stratification variable. DRSS response will be derived based on the imputed values. A total of 100 imputations will be used, followed by the MIANALYZE procedure to combine the estimates.

Secondary Efficacy Analysis:

Secondary endpoints involving a proportion:

Missing data for the secondary endpoints involving a proportion will be imputed using the same approach as described for the primary efficacy analysis (ICE handling and missing data imputation).

MAR imputation for the secondary endpoints involving a proportion will be generated based on a regression model with covariates for treatment group and the randomization stratification variable. Where appropriate, the dichotomous variable will be derived based on the imputed data.

Change from baseline secondary endpoints:

Handling of missing data of change from baseline in participants without ICEs:

- Multiple imputation of the value at Week 24 (for BCVA and CST) assuming missing at random (MAR).

Handling of missing data of change from baseline after ICEs:

- Single imputation using the worst value at each visit observed among all participants when applying composite variable strategy.
- Multiple imputation assuming MAR when applying treatment policy or hypothetical strategy.

Handling of data observed after ICEs:

- The value will be replaced by a single imputation using the observed worst value at each visit among all participants when applying composite variable strategy.
- Replaced by MAR multiple imputation algorithm based on the assumption that the intercurrent event would not have occurred when applying hypothetical strategy.
- Kept in the analysis when applying treatment policy strategy.

MAR imputation for the continuous secondary endpoints (BCVA and CST) will be generated based on a linear regression model adjusting for treatment group, the randomization strata, and BL score of the analyzed variable where applicable and then the change from baseline will be calculated. A total of 100 imputations will be used followed by the MIANALYZE procedure to combine the estimates.

Time to event secondary endpoints:

- Missing data for the time to event of secondary endpoints will be considered as censored. Participants who receive rescue medication will be censored at the start of rescue medication.

4.10.1.3 Multiple Comparisons/Multiplicity

No multiplicity adjustment is planned for this study.

4.10.1.4 Examination of Subgroups

The uniformity of the treatment effect for the primary efficacy variable will be examined for the following subgroups:

- Age (<65 years, ≥65 years)
- Gender (Male, Female)
- Race (White, Black or African American, Asian, Other)
- Ethnicity (Hispanic, Not Hispanic)
- BMI (Normal [<25 kg/m^2], Overweight [≥25 to <30 kg/m^2], Obese [≥30 kg/m^2])
- HbA1C (≤7%, >7%)
- Diagnosis of systemic hypertension (Y/N)
- Duration of Diabetes (<median, ≥median)
- Randomization strata (Screening DRSS:47 or 53 or 61B)
- Presence of fluid in the central subfield at Baseline (Y/N)

A forest plot with overall treatment effect included at the top will be provided.

4.10.2 Primary Efficacy Analysis

The primary efficacy outcome measure of the study will be the proportion of participants who have improved by ⩾2 steps from BL DRSS at Week 24. DRSS score levels will be recoded into 12 steps, provided by WRC.

Comparison of the primary endpoint will be made between each dose group and the combined vehicle control group using Mantel-Haenszel (MH) test of the difference in two proportions stratified by the randomization stratification factor (BL DRSS score 47 or 53 or 61B). The estimated MH risk difference will be summarized along with the two-sided 90% CI using MH stratum weights and the Sato variance estimator.

The primary efficacy analysis will be performed on the ITT Analysis Set, while the PPS will be used as a supplementary analysis. ICEs and missing data will be handled as described in Section 4.10.1.2.

In order to assess the robustness of the study conclusions to the choice of imputation method, the following sensitivity analyses will be performed on the ITT Analysis Set and PPS for the primary endpoint:

- An observed case analysis where only observed responses will be used for analysis.

  Observed data after initiation of rescue medication will be set to missing.
- A LOCF analysis where missing post-baseline responses will be imputed using the LOCF procedure.

  Measurements after rescue will be imputed using the last observation prior to rescue treatment; a non-response will be imputed if no post-baseline measurements are available.

A supplementary analysis will be performed on the ITT Analysis Set and PPS for the primary endpoint to assess the time to improvement of ⩾2 steps improvement from BL in DRSS score. Analyses will be based on observed data, where participants who receive rescue medication will be censored at the start of rescue medication (unless improvement of ⩾2 steps improvement from BL in DRSS score is achieved prior to starting rescue medication). This analysis will be presented based on the Kaplan-Meier method, along with a 90% confidence interval. A stratified log-rank test adjusting for the randomization stratification factor will be performed, to compare the combined vehicle control with each treatment group. Additionally, a stratified Cox proportional hazards model adjusting for the randomization stratification factor will be employed to estimate the hazards ratio between the combined vehicle control and each treatment group.

The primary efficacy analysis will also be performed by Baseline subgroups in Section 4.10.1.4, to assess the consistency of the treatment effect.

An additional sensitivity analysis may be performed to assess the proportion of participants with an improvement of ⩾2 steps improvement from the Visit 2/Baseline visit in DRSS score if needed.

4.10.3 Secondary Efficacy Analysis

The secondary outcome measures are described in Section 3.2.3. All secondary endpoints analyses will be performed on the Intention-to-Treat Analysis Set. The following secondary endpoint analyses will be repeated on the PPS:

- Proportion of participants who develop ASNV determined by the investigator at Week 24.
- Proportion of participants who develop CI-DME at Week 24.
- Change in DRSS steps is defined as DR worsening or improving by 1, 2, or ⩾3 steps.
- Proportion of participants with mild PDR (DRSS score 61B) at baseline who regress to NPDR (DRSS score ⩽ 53) by Week 24
- Mean and median change in BCVA (ETDRS letters) from baseline to Week 24.
- Lines gained/lost in BCVA (± 5, 10, and 15 ETDRS letters) at Week 24.
- Mean and median change in CST from baseline to Week 24.
- Proportion of participants who met the objective rescue criteria as defined in protocol section 2.14.9.
- Time to meet the objective rescue therapy criteria.
- Proportion of participants who develop VTC at Week 24

Continuous variable and changes from baseline will be summarized at each visit by treatment using descriptive statistics. Categorical variables will be summarized at each visit by treatment using frequencies and percentages.

The secondary endpoints involving proportions of participants will be analyzed using the same approach (MH test) as described for the primary efficacy analysis and based on the ITT Analysis set.

The categorical secondary endpoints will be compared using a proportional odds model with covariates for treatment group and the randomization stratification factor. If the assumption of proportional odds is rejected, a different analysis approach, such as a rank-based ANCOVA, will be

considered. The rank-based ANCOVA will be a two-sided test and a two-sided p-value will be presented.

The time to event secondary endpoints will be defined as time to first development of an event and will be estimated based on the Kaplan-Meier method, along with a 90% confidence interval. Analyses will be based on observed data, where participants who receive rescue medication will be censored at the start of rescue medication. A stratified log-rank test adjusting for the randomization stratification factor will be performed, to compare the combined vehicle control with each treatment group. Additionally, a stratified Cox proportional hazards model adjusting for the randomization stratification factor will be employed to estimate the hazards ratio between the combined vehicle control and each treatment group.

Continuous secondary endpoints involving change from BL to Week 24 will be analyzed using an analysis of covariance (ANCOVA) model with BL measurement of the variable as a covariate and treatment group and the randomization stratification factor as fixed factors. As a sensitivity analysis of the continuous secondary endpoints involving change from BL to Week 24, a rank-based ANCOVA model will be employed. The pair-wise comparisons of each OTT166 dose group versus combined vehicle control will be done in the ANCOVA model. Two-sided 90% confidence intervals for the difference of each OTT166 dose group minus combined vehicle control will be calculated.

The specific reasons for meeting rescue criteria as described in protocol section 2.14.9 will also be summarized using number and percentage by treatment group.

4.10.4 Exploratory Analysis

Exploratory imaging outcome measures are described in Section 3.2.4.

- All exploratory endpoint analyses will be performed on the Intention-to-Treat Analysis Set and will be based on observed data, where observed data after the initiation of rescue medication will be set to missing unless specified otherwise. The following exploratory endpoint analyses will be repeated on the PPS:
- Change in MV from baseline to Week 24 for 1, 3, and 6 mm regions.
- Change in area of retinal non-perfusion from baseline to Week 24 seen on widefield FA.
- Proportion of participants with intact FAZ seen on widefield FA at Week 24.

The exploratory endpoints involving proportions of participants will be analyzed using the same approach (MH test) as described for the primary efficacy analysis if not specified otherwise.

The time to event exploratory endpoints will be estimated based on the Kaplan-Meier method, along with a 90% confidence interval. A stratified log-rank test adjusting for the randomization stratification factor will be performed, to compare the combined vehicle control with each treatment group. Additionally, a stratified Cox proportional hazards model adjusting for the randomization stratification factor will be employed to estimate the hazards ratio between the combined vehicle control and each treatment group.

Continuous exploratory endpoints involving change from BL will be analyzed using an ANCOVA model with BL measurement of the variable as a covariate and treatment group and the randomization stratification factor as fixed factors. The pair-wise comparisons of each OTT166 dose group versus combined vehicle control will be done in the ANCOVA model. Two-sided 90% confidence intervals for the difference of each OTT166 dose group minus combined vehicle control will be calculated.

Within participant treatment effects will also be assessed for the primary endpoint (DRSS step change of ≥ 2 steps) and key secondary endpoints (BCVA change, CST change, and development of CI-DME). Differences in outcomes from screening to Week 24 between eyes for participants in the two active treatment groups (OTT166 5 mg/day, OTT166 10 mg/day) will be compared using an appropriate paired data analysis: McNemar's test for dichotomous data and either a paired t-test or a Wilcoxon signed-rank test for continuous outcome data (changes from baseline). Data for eyes that were treated with anti-VEGF and/or laser will be imputed with LOCF from the time of starting treatment with anti-VEGF and/or laser. The earliest concomitant medication in the fellow eye with ATC4CD="SL01A" will be considered for anti-VEFG treatment in the fellow eye. Additionally, the first ocular concomitant procedure in the fellow eye will be considered for laser treatment. The ocular concomitant procedures will be reviewed and confirmed by the clinical team.

Associations between baseline covariates and compliance may be assessed graphically: monthly compliance may be presented as a series of boxplots for each baseline subgroup, by treatment group and overall, allowing for the observation of compliance trends within or between groups.

4.11 Safety Evaluation

All safety summaries and analyses will be based upon the Safety Analysis Set by treatment group (i.e., vehicle control doses (5% BID, 5% QID), vehicle control combined, active OTT166 doses (5% BID, 5% QID), and OTT166 combined). Safety summaries for age, gender, race and ethnicity subgroups will be provided. If any safety measurements are repeated at the same visit after dosing on Day 1, then the last (non-missing) value of any repeated measurements will be used in the descriptive statistics and in the calculation of changes from BL. No inferential statistics will be performed for safety variables.

4.11.1 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 24.1 or higher. Intensity and causality of AE will be evaluated by the investigator. The intensity of AE will be graded according to the criteria in Appendix 4 of the Protocol.

Incomplete start and end date of the AEs will be imputed as described in Section 4.2.1.

Participants with multiple incidents of events for a given preferred term (PT) and system organ class (SOC) will be counted only once. Similarly, if participants experience multiple incidents of events for a given PT and SOC, the worst intensity will be used in the summaries presenting intensity, and the worst causality to study drug will be used in the summaries presenting causality, respectively.

Any missing intensity will be queried for completion and assessed by the Investigator based on AEs and serious AEs, as defined in Appendix 4 of the Protocol.

Any missing causality will also be queried for completion, and any events that still have missing causality in the final data will be deemed as 'Related'.

Adverse events of special interest (AESI) are defined as:

- Hypersensitivity reactions. Symptoms of hypersensitivity reactions (allergic reactions) include:
  - Rash
  - Wheezing and difficulty breathing

- Dizziness and fainting
- Swelling around the mouth, throat or eyes
- A fast pulse
- Sweating

Treatment-emergent adverse events (TEAEs) are defined as any adverse events occurring (onset date/time) at the time of or after dosing on Day 1. Where the AE start date is missing, AEs will be assumed to be treatment-emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the adverse event started prior to the first dose of study treatment.

Adverse event summaries will be sorted in terms of decreasing frequency for SOC, and PT within SOC, in the overall group, and then alphabetically for SOC, and PT within SOC if there are any ties in frequency.

Analysis and reporting for adverse events will be based on TEAEs. An overall summary of TEAEs will include the number and percentage of participants in the following categories:

- Number and percentage of participants with TEAEs
- Number and percentage of participants with drug-related TEAEs
- Number and percentage of participants with TEAEs leading to discontinuation of study drug
- Number and percentage of participants with TEAEs leading to study discontinuation
- Number and percentage of participants with treatment-emergent serious adverse event (TESAEs)
- Number and percentage of participants with drug-related treatment-emergent serious adverse event (TESAEs)
- Number and percentage of participants with TESAEs leading to discontinuation of study drug
- Number and percentage of participants with TESAEs leading to study discontinuation
- Number and percentage of participants by intensity of TEAEs (mild, moderate, severe)
- Number and percentage of participants who experienced death from any cause
- Number and percentage of participants with AESIs

Summaries of TEAEs by treatment group, SOC, and PT will be presented for the following:

- All TEAEs
- Ocular TEAEs in the Study Eye
- Ocular TEAEs in the Fellow Eye
- Non-Ocular TEAEs
- Drug-related TEAEs
- TEAEs leading to discontinuation of study
- TEAEs leading to discontinuation of study drug

- All TESAEs
- Ocular TESAEs in the Study Eye
- Ocular TESAEs in the Fellow Eye
- Non-Ocular TESAEs
- Drug-related TESAEs
- TESAEs leading to discontinuation of study
- TESAEs leading to discontinuation of study drug
- TEAEs by intensity
- AESIs
- AESIs leading to discontinuation of study
- AESIs leading to discontinuation of study drug

By-participant listings of corresponding TEAE data will also be provided. Separate listings will be presented for TEAEs leading to discontinuation of study drug and treatment-emergent serious adverse events (TESAEs). Any AEs occurring between signing of ICF and dosing on Day 1 will be presented in a separate listing. All listings will be done by treatment group and participant, detailing verbatim, SOC, PT, start date, stop date (if resolved), intensity, seriousness, outcome, action taken with respect to study drug, and relationship to study drug. The AE onset will also be shown relative (in number of days) to the day of the first study drug administration.

4.11.2 Clinical Laboratory Evaluation

All laboratory values will be reported and summarized using standard international (SI) units.

Continuous test results for each laboratory parameter and changes from baseline will be summarized at each visit by treatment using descriptive statistics. Categorical test results for each parameter will be summarized at each visit by treatment using frequencies and percentages.

The abnormal values will be flagged with ‘L’ for values below the lower limit of the clinical reference range and ‘H’ for values above the upper limit of the clinical reference range and included in the listings. A shift from baseline to out-of-normal range at each post-baseline visit will be summarized by treatment. If a participant has both low and high out-of-normal post-baseline value, each low and high will be counted separately. By-participant listings of laboratory parameters will also be provided.

Urinalysis results will be listed by participant and timepoint including changes from baseline for numeric variables, flags for any measurements that are outside the reference ranges and any repeated/unscheduled assessments.

The results of pregnancy and HbA1c testing will be listed for each participant.

4.11.3 Vital Signs

Vital signs measurements will be listed by participant and timepoint including changes from BL and any repeated/unscheduled measurements. Descriptive statistics (n, mean, SD, minimum, median,

maximum) will be provided by treatment group and timepoint for both absolute values and changes from BL.

4.11.4 Physical Examination

A listing of physical examination data by participant will be provided.

4.11.5 Data and Safety Monitoring Board (DSMB)

No DSMB is planned.

4.12 Planned Interim Analyses

There are no interim analyses planned.

4.13 Determination of Sample Size

The study will enroll approximately 210 participants, 70 participants in each OTT166 treatment group and 35 participants per group into each matching vehicle control group and allowing for around 5% study attrition, this will ensure at least 66 participants per treatment group and combined vehicle control group.

For the primary endpoint of improvement by at least 2 steps from baseline on the DRSS at Week 24, completing 66 participants per group (active treatment BID, active treatment QID, combined vehicle control) through Week 24 will yield at least 90% power to demonstrate that the active treatment is superior to placebo using a 1-sided alpha of 0.05, provided that the true treatment effect (active treatment efficacy rate minus control efficacy rate) is at least 0.20 and the control efficacy rate is at most 0.06.

The sample size calculation was computed in nQuery 8.7 using a Chi square test for two proportions (continuity corrected).

4.14 Changes in the Conduct of the Study or Planned Analysis

- Analyses described in the SAP supersede the planned analyses in the study protocol. Similarly, planned analyses may be amended as a result of planned blinded data reviews. Changes will be finalized prior to the database lock.

5 REFERENCES

[1] Mantel, N. and Haenszel, W., 1959. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the national cancer institute, 22(4), pp.719-748.

[2] Sato, T., 1989. On the variance estimator for the Mantel-Haenszel risk difference. Biometrics, 45(4), pp.1323-1324.

[3] Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials E9(R1) Final version (Step 4), Adopted on 20 November 2019. (https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf)

[4] Staurenghi, G., Feltgen, N., Arnold, J.J., Katz, T.A., Metzig, C., Lu, C. and Holz, F.G., 2018. Impact of baseline Diabetic Retinopathy Severity Scale scores on visual outcomes in the VIVID-DME and VISTA-DME studies. British Journal of Ophthalmology, 102(7), pp.954-958.

[5] Friedman, S. M., Almukhtar, T. H., Baker, C. W., Glassman, A. R., Elman, M. J., Bressler, N. M., Maker, M.P., Jampol, L.M., Melia, M. & Diabetic Retinopathy Clinical Research Network. (2015).

Topical nepafenec in eyes with non-central diabetic macular edema. Retina (Philadelphia, Pa.), 35(5), 944.

6 APPENDIX

6.1 Schedule of Activities

Study Procedure	Screening Visit 1	Baseline (BL) Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8[a]
Week	**--**	**0**	**4**	**8**	**12**	**16**	**20**	**24**
Day (visit window)	**-21 to -1**	**1**	**29 ± 7 days**	**57 ± 7 days**	**85 ± 7 days**	**113 ± 7 days**	**141 ± 7 days**	**169 ± 7 days**
Screening/Baseline (BL):								
Informed consent	X							
Inclusion/exclusion	X	X						
Medical and ophthalmic history	X							
Demographics	X							
Randomization		X						
Interval Medical Review:								
Review of concomitant medications	X	X	X	X	X	X	X	X
Continue study treatment[b]		X	X	X	X	X	X	X
Ocular Assessments:								
Refraction and BCVA (ETDRS)	OU	OU	OU	OU	OU	OU	OU	OU
Slit lamp examination	OU	OU	OU	OU	OU	OU	OU	OU
Assessment of cornea[c]	OU	OU	OU	OU	OU	OU	OU	OU
Intraocular pressure[d]	OU	OU	OU	OU	OU	OU	OU	OU
Indirect ophthalmoscopy	OU	OU	OU	OU	OU	OU	OU	OU

4-wide or 7-field color fundus photography (mandatory)[e]	OU	OU	SE	SE	OU	SE	SE	OU
SD-OCT (widefield if available)	OU	OU	SE	SE	OU	SE	SE	OU
OCTA (if available)	OU							OU
Wide field Fluorescein Angiography[m]	OU							OU
Gonioscopy	OU							OU
Non-Ocular Assessments:								
Physical examination[f]	X							
Vital signs[g]	X	X	X	X	X	X	X	X
Adverse events[h]	X	X	X	X	X	X	X	X
IP adherence review[n]			X	X	X	X	X X	
Laboratory Testing:[i]								
Hematology, blood chemistry, urinalysis, and HbA1c[j,l]	X							X
Pregnancy test, women of childbearing potential[k]	serum	urine	urine	urine	urine	urine	urine	urine

Abbreviations: BCVA: best corrected visual acuity; BL: baseline; ETDRS: Early Treatment Diabetic Retinopathy Study; HbA1c: glycosylated hemoglobin; OCTA: optical coherence tomography – angiography; OU: both eyes; SD-OCT: spectral-domain optical coherence tomography; SE: study eye;

[a] Participants who are withdrawn from the study before Week 24/Visit 8 will be asked to return to the clinic to complete the Visit 8 assessments.

[b] Participants will continue to receive study treatment based on exam findings, AE review, and verification that rescue criteria are not met.

[c] Cornea health assessment following instillation of fluorescein stain, including epithelium and endothelium, for signs of toxicity.

[d] Intraocular pressure should be measured at approximately the same time of day from visit to visit, if possible.

[e] An FP and FA must be performed once a participant has been diagnosed with worse than mild PDR, ASNV, or new onset or worsening CI-DME in the study eye, and before rescue treatment is given.

[f] May be performed at Visit 1 or Visit 2, Including height, weight, and body temperature.

[g] Vital signs (blood pressure and heart rate) will be measured after the participant has been sitting for 5 minutes at the screening visit. From baseline through end of study only blood pressure and heart rate will be measured.

[h] AEs will be collected from the time the ICF is signed until early termination or the end of study visit. If a participant withdraws from the study, ongoing AEs will be followed to the end of study visit or until the participant withdraws consent.

[i] All samples collected for laboratory assessments will be obtained prior to administration of study drug.

[j] At visits at which FA is performed, urinalysis samples will be collected before FA in order to avoid false elevations in urine protein values.

[k] For women of childbearing potential, a negative screening serum pregnancy test at Visit 1 is required before randomization. All women of childbearing potential will have a urine pregnancy test at each treatment visit starting at Visit 2 (Day 1); a negative urine pregnancy test is required before treatment is continued (randomized treatment, rescue treatment).

[l] Refer to Appendix 3.3 in protocol for clinical laboratory test details.

[m] Screening Fluorescein angiography is valid for 3 months. Participants undergoing rescreen do not need IVFA repeated within 3 months of original angiogram.

[n] Cornea should be assessed at the beginning before dilation.

[o] Adherence to the dosing regimen should be monitored continuously and discussed with the participant at each visit. Sites are expected to intervene if there is suboptimal adherence or absence of data transfer and document the intervention in the patient dashboard.

6.2 ETDRS Final Diabetic Retinopathy Severity Scale (For Individual Eye)

Level	Severity	Definition
10	DR absent	Microaneurysms and other characteristics absent
12*	Non-DR abnormalities	
14*	DR questionable	14A HE definite; microaneurysms absent
		14B SE definite; microaneurysms absent
		14C IRMA definite; microaneurysms absent
		14Z Venous loops >D/1; microaneurysms absent
15*	DR questionable	Hemorrhage(s) definite; microaneurysms absent
20	Microaneurysms only	Microaneurysms definite, other characteristics absent
35†	Mild NPDR	35A Venous loops ≥ D/1
		35B SE, IRMA, or VB = Q
		35D HE ≥ D/1 (< STD 3 in any field)
		35E HE ≥ M/1 (≥ STD 3 in any field)
		35F SE ≥ D/1 (any definite)
43	Moderate NPDR	43A H/Ma = M/4-5 (> STD 1(4+ fields) -OR- ≥ STD 2A (1 field))
		43B IRMA = D1-3 (any definite < STD 8A)
47	Moderately Severe NPDR	47A Both L43 characteristics
		47B IRMA = D/4-5 (< STD 8A)
		47C H/Ma = S/2-3 (< STD 2B)
		47D VB = D/1
53	Severe NPDR	53A ≥ 2 of the L47 characteristics
		53B H/Ma ≥ S/4-5 (≥ STD 2A in 4+ fields or ≥ 2B in any field)
		53C IRMA ≥ M/1 (≥ STD 8A)
		53D VB ≥ D/2-3 or S1 (≥ STD 6 in 2+ fields or ≥ 6B in any field)
53E	Very Severe NPDR	53E ≥ 2 of 53B, 53C, and 53D
60	Inactive PDR	60 Scatter or Local Rx
61	Mild PDR	61A FPD and/or FPE only (regressed PDR)
		61B NVE < ½ DA in ≥ 1 field
65	Moderate PDR	65A NVE ≥ M/1 (≥ ½ DA in ≥ 1 field)
		65B NVD = D (< STD 10A) ; and VH and PRH = A or Q
		65C VH or PRH = D (< 1 DA) and NVE < M (< ½ DA) and NVD absent

71/75	High-risk PDR	71A VH or PRH ≥ M/1 (≥ 1 DA)
		71B NVE ≥ M/1 (≥ ½ DA 1 field) **and** VH or PRH ≥ D/1 (any)
		71C NVD = D (< STD 10A) **and** VH or PRH ≥ D/1 (any)
		71D NVD ≥ M (≥ STD 10A)
		75 NVD ≥ M (≥ STD 10A) **and** VH or PRH ≥ D/1 (any)
81	Advanced PDR: fundus partially obscured, center of macula attached	NVD = cannot grade, or NVD < D and NVE = cannot grade in ≥ 1 field and absent in all others; and retinal detachment at center of macula < D
85	Advanced PDR: posterior fundus obscured, or center of macula detached	85A VH = VS (obscuring) in Field 1 or 2 85B Retinal detachment at center of macula = D (present)
90	Cannot grade, even for 81 or 85	

Abbreviations: DR = diabetic retinopathy; NPDR = nonproliferative DR; PDR = proliferative DR; HE = hard exudates; SE = soft exudates; IRMA = intraretinal microvascular abnormalities; VB = venous beading; H/Ma = hemorrhages/microaneurysms; NVE = new vessels elsewhere; NVD = new vessels on or adjacent to optic disc; VH = vitreous hemorrhage; PRH = preretinal hemorrhage

Severity categories are of the form (maximum severity/extent), where maximum severity can be absent (A), questionable (Q), definitely present (D), moderate (M), severe (S), or very severe (VS), and extent is the number of photographic fields at that severity level. For example, M/2-3 means there are two or three fields from fields 3 to 7 with moderate severity, and none with higher severity.

* Levels 12, 14, and 15 are not considered separate steps in the scale, but are pooled with level 10 or 20 (or excluded).

† NPDR levels 35 and above all require presence of microaneurysms.

6.3 AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN (SAP)

Rationale for the amendment

This statistical analysis plan has been amended to incorporate changes made in the current protocol amendment v5.0 of 10 October 2023 as well as additional changes not in the current protocol.

Changes below that are not shown in the current version of the protocol should supersede any analysis described in the current protocol v5.0.

6.3.1 Amendment 1

Change # 1
List of Abbreviations
The following abbreviations have been added or updated:

- **IRMA: Intraretinal microvascular abnormality was added**
- TMV: Total Macular Volume has been changed to **MV: Macular Volume**

Change # 2

Section 2 STUDY OBJECTIVES AND ENDPOINTS

- **Section 2.1, 2.2, and 2.3** have been combined as Section 2 and updated to present all the study objectives and endpoints in a table format for clarity
- The following endpoint has been updated (the text in bold has been added):
  - Proportion of participants who develop ASNV **determined by the investigator** at Week 24
- The following secondary endpoint has been deleted from this section and throughout the SAP
  - **Mean and median step change in DRSS step from baseline to Week 24**

Change# 3
Section 3.1 Overall Study Design and Plan:

- The following changes in bold have been made in this section
  - The study population will be participants with Type 1 or Type 2 diabetes under adequate control as evidenced by glycosylated hemoglobin **(HbA1c⩽12.0%),** who show typical fundus alterations and have been diagnosed with moderately severe to severe NPDR or mild PDR, as defined by the DRSS Steps 47, 53 and 61B (Appendix 6.2). Participants may have fluid in the macula with BCVA of ⩾69 Early Treatment Diabetic Retinopathy Study (ETDRS) letters (Snellen equivalent 20/40) and Central Subfield Thickness (CST) of ⩽ **325** µm, **and a normal foveal contour, which needs to be confirmed by an independent central reading center, in at least 1 eye**.

Change# 4
Section 3.2.1 Primary Endpoint – Safety

- The Primary Endpoint - Safety section has been updated to only include the primary endpoint for safety shown in bold below and remove all other secondary safety endpoints:
  - **The primary safety outcome measure of the study will be the proportion of participants who develop TEAEs through Week 24.**

Change# 5
Section 3.2.2 Primary Efficacy Endpoint and Estimand

- The following changes (text shown in bold below) have been made in this section:
  - The primary clinical question of interest is to determine the effectiveness of different doses of OCT166 versus vehicle control on ⩾2 steps improvement from BL in DRSS score at Week 24 **without requiring rescue treatment if the study drug is not discontinued prior to Week 24**. This will support the dose selection for subsequent pivotal studies of OTT166. The estimand for the primary efficacy analysis is defined as follows,
- Participant population attribute of the primary Estimand has been updated:
  - DR participants with **Type 1 or Type 2 diabetes under adequate control (HbA1c ⩽12.0%) who show typical fundus alterations and have been diagnosed with moderately severe to severe NPDR or mild PDR as defined by the DRSS Steps 47, 53, and 61B, with BCVA ⩾ 69 ETDRS letters, CST of ⩽ 325 µm, and a normal foveal contour.**
- The following ICE strategies have been changed:
- ICE b) has been changed from "Death due to DR" to "**Death**", and strategy has been changed from Composite strategy (non-response) to **Hypothetical**.

- ICE c) Strategy has been changed from Hypothetical to **Composite strategy (non-response).**
- ICE e) Strategy has been changed from Hypothetical to **Treatment policy**.
- ICE e) and f) compliance evaluation time range has been changed from every 4 weeks **through Week 24**
- ICE f) upper bound of compliance has been changed from 100% to **90%**

- The following paragraph (in bold) has been moved from this Section to Section 4.10.1.2:
  - **If a participant experiences more than one ICE, ICE(s) with a composite variable strategy take priority over ICEs with other strategies. And ICEs with a hypothetical strategy take priority over ICEs with treatment policy strategy.**

- The Supplemental Estimand description was moved from Section 4.10.1.2 to this Section and the strategy was changed from treatment policy for ICEs c) through f) to having ICEs c) and e) use Hypothetical strategy as described below in bold:
  - **A Supplemental Estimand will be defined and analyzed where the ICEs defined above will be handled using alternative strategies as follows: c, and e would be changed to Hypothetical strategy**.

Change# 6
Section 3.2.3 Secondary Endpoints

- The following secondary endpoint has been deleted from this section:
  - **Mean and median change in DRSS step from baseline to Week 24.**
- The following changes (text in bold) have been made to the following secondary endpoints:
  - Proportion of participants who develop ASNV **determined by the investigator** at Week 24.
  - Proportion of participants who develop CI-DME at Week 24.
    - CI-DME is defined as the presence of any fluid within **the central subfield in participants who have no fluid at baseline or CST> 325 µm**.
  - Proportion of participants who met the **objective** rescue criteria as defined in protocol section 2.14.9.
  - Time to meet **the objective** rescue therapy criteria as defined in protocol section 2.14.9

Change # 7
Section 3.2.4 Exploratory Endpoints

The following changes in bold have been made:

- Change in Total Macular Volume (TMV) from baseline to Week 24 for 1, 3, and 6 mm regions endpoint has been updated to Change in Macular Volume (MV) and moved from the bottom of this section to the top; the definition has also been updated (see in bold below):
  - **Change in Macular Volume (MV) from baseline to Week 24 for 1, 3, and 6 mm regions.**
    - **MV for the 1 mm region is the central 1 subfield, for the 3 mm region is the middle 5 subfields and for the 6 mm region is all 9 subfields as provided by the imaging center**
- The following endpoints have been deleted:

- **Proportion of participants who do not have non-CI-DME at baseline that develop non-CI-DME at Week 24.**
  - **Non-CI-DME is defined as central subfield thickness below the threshold defined in Friedman et. al. (2015) for Zeiss Cirrus and meeting one of the following: 1) At least 2 non-central subfields with thickness values above threshold; or 2) At least 1 non-central subfield with thickness value more than 15 µm above threshold.**
- **Time to development of non-CI-DME (in those participants who do not have non-CI-DME at baseline).**
- **Time to resolve non-CI-DME (in those participants who do have non-CI-DME at baseline).**
- **Optical Coherence Tomography (OCT) endpoints:**
  - **Proportion of participants with appearance or disappearance of subretinal fluid or cysts in the 9 subfields at Week 24**

Change# 8

Section 4.4.1 Analysis Sets

- The following changes in bold have been made to the definition of the ITT Analysis Set:
  - All randomized participants **who received at least one dose of the study drug**. Participants will be included into the analysis as randomized. **This corresponds to what is referred to in ICH E9 as the Full Analysis Set.**
- FAS definition and references throughout the SAP have been deleted
- The following changes in bold have been made to this section:
  - **The** ITT **Analysis Set** will be used for all efficacy **analyses** and BL analyses **except participant disposition which will be shown for the ITT Analysis Set and for the Safety Analysis Set if the two analysis sets differ.** The PP set will be determined prior to database lock and will be used as a supplementary analysis of the primary endpoint.
  - The number and percentage of participants in each analysis set by treatment group will be summarized for the **ES**. **Inclusion in the PP set** will be **indicated in the listing of participant disposition for the ITT Analysis Set**.
  - **Unless otherwise specified, participant disposition, demographics, and baseline characteristics will be summarized by the randomized treatment group (i.e., vehicle control doses (5% BID, 5% QID), vehicle control combined, active OTT166 doses (5% BID, 5% QID)), and overall**.

Change # 9

Section 4.4.2 Disposition of Participants

- The following change in bold has been made:
  - Participant disposition will be summarized **for the ITT Analysis Set** by treatment and overall, for:
- The following participant disposition characteristics in bold have been deleted:
  - **Number of participants screened**
  - **Number of participants with screen failure**
- The following paragraph in bold has been added:

- **Participant disposition for participants who were screen failures (defined as participants in the Enrolled Set who were not randomized and are not included in the ITT Analysis Set) will be summarized and will include the reason for screen failure.**

Change# 10
Section 4.4.3 Protocol Deviations

- The following changes in bold have been made:
  - The number and percentage of participants with at least one major protocol deviation as well as major protocol deviation categories **for the ITT Analysis Set** by treatment and overall will be summarized.

Change# 10
Section 4.5 Demographics and Baseline Characteristics

- The following changes in bold have been made:
  - Central **Subfiel**d Thickness **(CST)** at Baseline
  - HbA1C **(≤7%, >7%)**
- The following subgroup has been added:
  - **Presence of fluid in the central subfield at Baseline (Y/N)**

Change# 11
Section 4.8 Rescue Therapy

- The following changes in bold have been made in this section:
  - Participants meeting one or all the following criteria would qualify for rescue with anti-VEGF therapy and/or PRP laser after consultation with the Medical Monitor:
    Objective Rescue Therapy Criteria:
    - Worsening DRSS level to 65 or higher as determined by the central reading center
    - Decrease in BCVA of ⩾ 10 ETDRS letters from BL associated **with presence of new fluid or an increase in fluid in the central subfield compared to baseline** at any visit
    - Decrease in BCVA of 5 to 9 ETDRS letters from BL at two consecutive visits 4 weeks apart associated **with presence of new fluid or an increase in fluid in the central subfield compared to baseline**
    - Increase in CST of ⩾ 75 microns from BL associated **with presence of new fluid or an increase in fluid in the central subfield compared to baseline**

    Subjective Rescue Therapy Criteria:
    - Any **other** condition that, in the clinical opinion of the investigator, warrants rescue therapy after consultation with the Medical Monitor

Change# 12
Section 6.1.1 Treatment Exposure

The following changes in bold have been made in this section:

- Exposure to study drug will be summarized using **the ITT Analysis Set and also** the Safety Analysis Set. The total number of **doses** administered, and the duration of treatment will be summarized in a summary table by treatment group and overall.

Change# 13
Section 4.9.2 Treatment Compliance

- The following changes and monthly compliance in bold have been added:
  - A summary of the treatment compliance will be provided **for the ITT Analysis Set and also the Safety Analysis Set** by treatment and overall.
  - **Monthly (every 4 weeks) treatment compliance (%) will be calculated as:**
    - $$\frac{\textbf{Total Number of Doses of Study Drug Administered each Month}}{\textbf{Total Number of Doses of Study Drug Planned for each month}}$$

Change# 14
Section 4.10.1 Analysis and Data Conventions
The following changes in bold have been made in this section:

- This study is designed to evaluate the superiority of two doses of OTT166 (5 mg/day and 10 mg/day) compared to vehicle control for the treatment of DR. For each dose within each endpoint, the null hypothesis is that there is no **difference** between OTT166 and vehicle control **or the difference is in the direction of inferiority of OTT166**; the alternative hypothesis is that there is a difference between OTT166 and vehicle control in the direction of superiority of OTT166.

Change# 15
Section 4.10.1.2 Handling of Dropouts or Missing Data

- Section was renamed to **Handling of Missing Data and Data After Intercurrent Events**
- Supplemental estimand description was moved to section **3.2.2** above
- BL HbA1C has been deleted from the MAR imputation model description in this section
- **Primary Efficacy Analysis**: was added as the subheading just under the section heading to distinguish methods used for the primary efficacy analysis compared to the secondary efficacy analysis which also had a subheading added

The following changes in bold have been made in this section:

- Handling of DRSS response at Week 24 data missing after ICEs:
  - Single imputation (non-response) when applying composite variable strategy.
  - Multiple imputation assuming MAR when applying hypothetical and treatment policy strategy. **For hypothetical strategy, data up until the ICE occurred will be used in the imputation; for treatment policy strategy, all available data will be used in the imputation**.

The following paragraph has been moved from Section **3.2.2** to this section:

- **If a participant experiences more than one ICE, ICE(s) with a composite variable strategy take priority over ICEs with hypothetical strategies, which take priority over ICEs with a treatment policy strategy**

The following changes in bold have been made in this section:

MAR imputation will be generated based on a logistic regression model for DRSS response (Y/N) with covariates for **randomized** treatment group **and** the randomization stratification variable. A total of 100 imputations will be used, followed by **the** MIANALYZE procedure to combine the estimates.

Secondary Efficacy Analysis:

Secondary endpoints involving a proportion:

Missing data for the secondary endpoints **involving a proportion** will be imputed using the same approach as described for the primary efficacy analysis (ICE handling and missing data imputation).
MAR imputation for the secondary endpoints involving a proportion will be generated based on a regression model with covariates for treatment group **and** the randomization stratification variable.

***Change from baseline secondary endpoints**:*

Handling of missing data of change from baseline in participants without ICEs:

- Multiple imputation **of the value at Week 24 (for BCVA and CST)** assuming missing at random (MAR).

Handling of missing data of change from baseline after ICEs:

- Single imputation using **the worst** value **at Week 24 observed among all participants** when applying composite variable strategy.
- Multiple imputation assuming MAR when applying treatment policy or hypothetical strategy.

Handling of data observed after ICEs:

- The value will be replaced by a single imputation using **the worst** value **at Week 24 observed among all participants** when applying composite variable strategy.
- Replaced by MAR multiple imputation algorithm based on the assumption that the intercurrent event would not have occurred when applying hypothetical strategy.
- Kept in the analysis when applying treatment policy strategy.

MAR imputation for the continuous secondary endpoints **(BCVA and CST)** will be generated based on a linear regression model adjusting for treatment group, the randomization strata, and BL score **of the analyzed variable** where applicable, **and then the change from baseline will be calculated.** A total of 100 imputations will be used followed by **the** MIANALYZE procedure to combine the estimates.

***Time to event secondary endpoints**:*

- Missing data for the time to event **of** secondary endpoints will be considered as censored.

Change# 16
Section 4.10.1.4 Examination of Subgroups
The following changes in bold has been added:

- Age (<65 years, ≥65 years)
- Gender **(Male, Female)**
- Race **(White, Black or African American, Asian, Other)**
- Ethnicity **(Hispanic, Not Hispanic)**
- BMI (Normal [<25 kg/m^2], Overweight [≥25 to <30 kg/m^2], Obese [≥30 kg/m^2])
- Diagnosis of systemic hypertension **(Y/N)**
- HbA1C **(⩽7%, >7%)**

The following subgroup in bold has been added:

- **Presence of fluid in the central subfield at Baseline (Y/N)**

Change# 17
Section 4.10.2 Primary Efficacy Analysis
The following changes in bold have been made:

- The primary efficacy analysis will be performed on the ITT **Analysis Set**, while the PPS will be used as a supplementary analysis. ICEs and missing data will be handled as described in Section 4.10.1.2.
- **In order to assess the robustness of the study conclusions to the choice of imputation method, the following sensitivity analyses will be performed for the primary endpoint:**
    - **An observed case analysis where only observed responses will be used for analysis**
    - **A last observation carried forward (LOCF) analysis where missing post-baseline responses will be imputed using the LOCF procedure**
    - **Measurements after rescue will be imputed using the last observation prior to rescue treatment; a non-response will be imputed if no post-baseline measurements are available.**

Change# 18
Section 4.10.3 Secondary Efficacy Analysis
The following changes have been made in this section:
The specific reasons for meeting rescue criteria as described in protocol section 2.14.9 will also be **summarized using number and percentage by treatment group**.

Change# 19
Section 4.10.4 Exploratory Analysis
The following paragraphs in bold have been deleted from this section:

- **The proportion of participants who have non-CI-DME at baseline that is resolved (no non-CI-DME) by Week 24 and the proportion of participants who do not have non-CI-DME at baseline that develop non-CI-DME by Week 24 will be summarized using a shift table.**
- **The normality assumption will be visualized based on the histogram of the residuals, and if the data are skewed, then a rank-based ANCOVA will be performed**.

The following changes in bold have been made in this section:

- Within participant treatment effects will also be assessed **for the primary endpoint (DRSS step change of ⩾ 2 steps) and key secondary endpoints (BCVA change, CST change, and development of CI-DME).**

The following paragraph in bold has been added to this section:

- **Associations between baseline covariates and compliance will be assessed graphically: monthly compliance will be presented as a series of boxplots for each baseline subgroup, by treatment group and overall, allowing for the observation of compliance trends within or between groups.**

Change# 20
Section 4.11 Safety Evaluation
The following changes in bold have been made in this paragraph:

- All safety summaries and analyses will be based upon the Safety Analysis Set by treatment **group (i.e., vehicle control doses (5% BID, 5% QID), vehicle control combined, active OTT166 doses (5% BID, 5% QID), and OTT166 combined**. Safety summaries for age, gender, race and ethnicity subgroups will be provided. If any safety measurements are repeated after dosing on Day 1, then the last (non-missing) value of any repeated measurements will be used in the descriptive statistics and in the calculation of changes from BL. No inferential statistics will be performed for safety variables.

Change# 21
Section 4.11.1 Adverse Events
The following AESI definition has been updated to match the definition in the amended protocol section:

- Adverse events of special interest (AESI) are defined as:
  Hypersensitivity reactions. Symptoms of hypersensitivity reactions (allergic reactions) include:
  - **Rash**
  - **Wheezing and difficulty breathing**
  - **Dizziness and fainting**
  - **Swelling around the mouth, throat or eyes**
  - **A fast pulse**
  - **Sweating**

Number and percentage of participants with TEAEs leading to death has been deleted and replaced with the **Number and percentage of participants who experienced death from any cause**

The following analyses in bold have been added:

- **Number and percentage of participants with TEAEs leading to discontinuation of study**
- **Number and percentage of participants with TESAEs leading to discontinuation of study**
- **Number and percentage of participants with AESIs leading to discontinuation of study drug**

Change# 22
Section 4.14 Changes in the Conduct of the Study or Planned Analysis
The following changes in bold have been made in this section:

- **Analyses described in the SAP supersede** the planned analyses in the study protocol. Similarly, planned analyses may be amended as a result of planned blinded data reviews. Changes will be finalized prior to the database lock.

Change# 23
Appendix 6.2 ETDRS Final Diabetic Retinopathy Severity Scale (For Individual Eye)

- ETDRS Final Diabetic Retinopathy Severity Scale has been updated with a new table that shows the subcategories and definitions within each DRSS level.

6.3.2 Amendment 2

Rationale for the amendment

This statistical analysis plan has been amended to incorporate changes made in the current protocol amendment v5.0 of 10 October 2023 as well as additional changes not in the current protocol.

Change # 1
Section 1 INTRODUCTION
The protocol version was updated:

- Study Protocol Version 5.0 dated October 10, 2023.

Change # 2
Section 2 STUDY OBJECTIVES AND ENDPOINTS

- The following objectives and endpoints have been updated (the text in bold has been added):

Secondary Objectives:	**Secondary Endpoints:**
To determine if topical OTT166 will prevent or delay the occurrence of visually threatening complications (VTC), defined as PDR worse than mild and/or anterior segment neovascularization (ASNV) determined by the investigator **and/or CI-DME (defined below)**	• Proportion of participants developing worse than mild PDR (DRSS 65 and above) at Week 24 • Time to development of PDR worse than mild (DRSS 65 and above) • Proportion of participants who develop ASNV determined by the investigator at Week 24 • **Proportion of participants who develop VTC at Week 24** • **Time to development of PDR worse than mild (DRSS 65 and above) or CI-DME**
To determine if topical OTT166 will prevent or delay the occurrence of CI-DME (CI-DME is defined as the presence of fluid in the central subfield in participants who have no fluid at baseline or CST> 325 µm)	• Proportion of participants who develop CI-DME at Week 24 • Time to development of CI-DME

Change# 3
Section 3.1 Overall Study Design and Plan:

- The following changes in bold have been made in this section
  - Approximately 210 participants diagnosed with moderately severe to severe NPDR or mild PDR and who are treatment naïve (*i.e.*, no prior anti-Vascular Endothelial Growth Factor [VEGF] or laser (focal, grid, Pan-Retinal Photocoagulation [PRP]) administered) will be randomized 2:2:1:1 into the following groups: OTT166 5% BID, OTT166 5% QID, vehicle control 5% BID, vehicle control 5% QID. Randomization will be stratified by **Screening** DRSS score (47 or 53 or 61B).

Change# 4

Section 3.2.2 Primary Efficacy Endpoint and Estimand

- The Primary Estimand table has been updated to clarify the intercurrent events and intercurrent event handling (changes in bold):

	Primary Estimand
Treatment conditions of interest	OTT166 5 mg/day, OTT166 10 mg/day and matching vehicle control
Participant Population	DR participants with Type 1 or Type 2 diabetes under adequate control (HbA1c ⩽12.0%) who show typical fundus alterations and have been diagnosed with moderately severe to severe NPDR or mild PDR as defined by the DRSS Steps 47, 53, and 61B, with BCVA ⩾ 69 ETDRS letters, CST of ⩽ 325 µm, and a normal foveal contour.
Endpoint	≥2 steps improvement from BL in DRSS score at Week 24
Population level summary	Difference in proportion of participants achieving ≥2 steps improvement from BL in DRSS score at Week 24 between each OTT166 treatment group and the combined vehicle control group.
Intercurrent events (ICEs) and strategies to handle ICEs	g) Rescue treatment **in the study eye** due to worsening DRSS level to 65 or higher prior to assessment of DRSS at Week 24 *Composite strategy (non-response)* h) Death prior to DRSS assessment at Week 24 *Hypothetical strategy* i) Use of rescue treatment **in the study eye** for any other reason prior to DRSS assessment at Week 24 *Composite strategy (non-response)* j) Premature discontinuation of study drug for any reason prior to DRSS assessment at Week 24 *Hypothetical strategy* k) Study drug compliance <75% through Week 24 *Treatment policy strategy* l) Study drug compliance between 75% to <90% through Week 24 *Treatment policy strategy*
	ICE(s) with a composite variable strategy take priority over ICEs with other strategies. And ICEs with a hypothetical strategy take priority over ICEs with treatment policy strategy.

Change# 5

Section 3.2.3 Secondary Endpoints

- The following endpoints were added:
  - Proportion of participants who develop VTC at Week 24
  - Time to development of PDR worse than mild (DRSS 65 and above) or CI-DME

Change# 6

Section 4.2 General Presentation Considerations

- The Baseline definition was updated (changes in bold):
  - 'Baseline' is defined as the last available pre-treatment assessment **for safety assessments. For DRSS score Baseline data, the Baseline value is defined as the Screening/Visit 1 result. Pre-treatment values for DRSS were collected at the Screening/Visit 1 using fluorescein angiography in addition to fundus photography whereas Baseline/Visit 2 used fundus photography only. Participants were enrolled into the study and randomized to the study treatment based on the Screening/Visit 1 DRSS result. All other efficacy analyses that do not use DRSS score will use the last available pre-treatment assessment as Baseline. This is considered acceptable as this measurement is still the best representation of the Baseline value of the given assessment since it is highly unlikely that the study drug could have an impact on any measurement in a short period of time.**
- The following sentence was added:
  - Time to event analyses will be calculated from Study Day 1, including time to DRSS-related endpoints comparing to the Baseline value. These analyses will consider only post-Baseline assessments. The Visit 2/Baseline visit will not be considered.

Change # 7

Section 4.4.1 Analysis Sets

The following changes in bold have been made:

- The ITT Analysis Set will be used for all efficacy analyses and BL **data** analyses except participant disposition which will be shown for the ITT Analysis Set and for the Safety Analysis Set if the two analysis sets differ. **All randomized participants are expected to be dosed and any exceptions will be described in the CSR.** The PP set will be determined prior to database lock and will be used as a supplementary analysis of the primary endpoint.
  For analyses based on the SS, participants randomized to a control treatment group who receive any dose of active OTT166 treatment will be analyzed in the corresponding OTT166 treatment group (BID or QID).

Change# 8

Section 4.4.3 Protocol Deviations

- The following sentence was added:
  - This summary will also include a summary of participants excluded from the PPS due to major protocol deviations by category.

Change # 9

Section 4.5 Demographics and Baseline Characteristics

- The following change in bold has been made:
  - DRSS steps at **the** Baseline **Visit**

Change# 10

Section 4.8 Rescue Therapy

- The following change in bold has been made:
  - Participants meeting one or all the following criteria **in the study eye** would qualify for rescue with anti-VEGF therapy and/or PRP laser after consultation with the Medical Monitor
- The following objective rescue criteria was added:
  - Development of anterior segment neovascularization (ASNV)
- The following text was added:
  - The start date of rescue therapy in the study eye will be determined by the earliest concomitant medication in the study eye with ATC4CD="SL01A" or the earliest ocular concomitant procedure confirmed by the clinical team to be rescue therapy. This will be subset on the participants who have completed the "Rescue Therapy for Ocular Diabetic Complications" eCRF.

Change# 11
Section 4.9.2 Treatment Compliance

- The following text has been added (additions in bold):

  The total number of doses of study drug administered is the actual number of drops administered. In the event a participant takes study drug after their EOS date, the total number of doses administered will not include any doses recorded after the EOS date.

  The total number of doses of study drug planned is calculated based on a participant's duration in the study and the randomized treatment. A participant randomized to BID treatment will have 2 times the (date of EOS – date of first dose + 1). A participant randomized to QID treatment will have 4 times the (date of EOS – date of first dose + 1).

  For participants who take rescue therapy in the study eye during the study, the total number of doses of study drug planned is calculated based on a participant's treatment duration prior to rescue and randomized treatment. A participant randomized to BID treatment will have 2 times the (date of EOT – date of first dose + 1). A participant randomized to QID treatment will have 4 times the (date of EOT – date of first dose + 1).

  A summary of the treatment compliance will be provided for the ITT Analysis Set and also the Safety Analysis Set by treatment and overall.

  Monthly (every 4 weeks) treatment compliance (%) will be calculated as:

$$\frac{\text{Total Number of Doses of Study Drug Administered each Month}}{\text{Total Number of Doses of Study Drug Planned for each month}} \times 100\%$$

  The total number of doses of study drug planned is calculated based on a participant's duration in the study and the randomized treatment. A participant randomized to BID treatment will have 2 times the minimum of 28 or the (date of EOS – date of first dose in month + 1). A participant randomized to QID treatment will have 4 times the minimum of 28 or the (date of EOS – date of first dose in month + 1).

  For participants who take rescue therapy in the study eye during the study, the total number of doses of study drug planned is calculated based on a participant's treatment duration prior to rescue and randomized treatment. A participant randomized to BID

treatment will have 2 times the minimum of 28 or the (date of EOT – date of first dose in month + 1). A participant randomized to QID treatment will have 4 times the minimum of 28 or the (date of EOT – date of first dose in month + 1).

Number and percentage of participants with compliance in categories (<75%, 75-<90%, ≥90%) will be provided for overall and monthly treatment compliance.

Change# 12
Section 4.10.1 Analysis and Data Conventions

- The following changes in bold have been made in this section:
  - For each dose within each endpoint, the null hypothesis is that there is no difference between OTT166 and vehicle control or the difference is in the direction of inferiority of OTT166, **and** the alternative hypothesis is that there is a difference between OTT166 and vehicle control in the direction of superiority of OTT166.
  - The vehicle control group will be combined for all efficacy analyses and two-sided 90% confidence intervals will be presented. **A comparison between the two control groups will be performed for the primary endpoint to confirm the groups are comparable.**

Change# 13
Section 4.10.1.1 Adjustments for Covariates
The following changes in bold have been made in this section:

- The primary efficacy analysis will be adjusted for the randomization stratification factor, the **Screening** DRSS score (47 or 53 or 61B).

Change# 14
Section 4.10.1.2 Handling of Missing Data and Data After Intercurrent Events

- The following changes have been made (changes in bold):
  - Primary Efficacy Analysis: MAR imputation will be generated based on a ~~logistic~~ regression model for DRSS ~~response (Y/N)~~ **levels** with covariates for randomized treatment group and the randomization stratification variable. **DRSS response will be derived based on the imputed values.** A total of 100 imputations will be used, followed by the MIANALYZE procedure to combine the estimates.
  - Secondary Efficacy Analysis:
    - Secondary endpoints involving a proportion: MAR imputation for the secondary endpoints involving a proportion will be generated based on a l~~ogistic~~ regression model with covariates for treatment group and the randomization stratification variable. **Where appropriate, the dichotomous variable will be derived based on the imputed data.**
    - Change from baseline secondary endpoints:

      Handling of missing data of change from baseline after ICEs:

      - Single imputation using the worst value at **each visit** ~~Week 24~~ observed among all participants when applying composite variable strategy.
      - Multiple imputation assuming MAR when applying treatment policy or hypothetical strategy.

      Handling of data observed after ICEs:

- The value will be replaced by a single imputation using the **observed** worst value at **each visit** ~~Week 24~~ among all participants when applying composite variable strategy.
- Time to Event secondary endpoints: Missing data for the time to event of secondary endpoints will be considered as censored. **Participants who receive rescue medication will be censored at the start of rescue medication.**

Change# 15
Section 4.10.1.4 Examination of Subgroups
The following subgroup was updated:

- ~~Baseline~~ **Randomization strata (Screening** DRSS:47 or 53 or 61B)

Change# 16
Section 4.10.2 Primary Efficacy Analysis

- The following text was added (in bold)
  - In order to assess the robustness of the study conclusions to the choice of imputation method, the following sensitivity analyses will be performed on the **ITT Analysis Set and PPS** for the primary endpoint:
    - An observed case analysis where only observed responses will be used for analysis. **Observed data after initiation of rescue medication will be set to missing.**
  - A supplementary analysis will be performed **on the ITT Analysis Set and PPS** for the primary endpoint to assess the time to improvement of ≥2 steps improvement from BL in DRSS score. **Analyses will be based on observed data, where participants who receive rescue medication will be censored at the start of rescue medication (unless improvement of ≥2 steps improvement from BL in DRSS score is achieved prior to starting rescue medication).**
  - **An additional sensitivity analysis may be performed to assess the proportion of participants with an improvement of ⩾2 steps improvement from the Visit 2/Baseline visit in DRSS score if needed.**

Change# 17
Section 4.10.3 Secondary Efficacy Analysis
The following text has been added:
The following secondary endpoint analyses will be repeated on the PPS:

- Proportion of participants who develop ASNV determined by the investigator at Week 24.
- Proportion of participants who develop CI-DME at Week 24.
- Change in DRSS steps is defined as DR worsening or improving by 1, 2, or ⩾3 steps.
- Proportion of participants with mild PDR (DRSS score 61B) at baseline who regress to NPDR (DRSS score ⩽ 53) by Week 24
- Mean and median change in BCVA (ETDRS letters) from baseline to Week 24.
- Lines gained/lost in BCVA (± 5, 10, and 15 ETDRS letters) at Week 24.

- Mean and median change in CST from baseline to Week 24.
- Proportion of participants who met the objective rescue criteria as defined in protocol section 2.14.9.
- Time to meet the objective rescue therapy criteria.
- Proportion of participants who develop VTC at Week 24

The following text in bold has been added:

The categorical secondary endpoints will be compared using a proportional odds model with covariates for treatment group and the randomization stratification factor. If the assumption of proportional odds is rejected, a different analysis approach, such as a rank-based ANCOVA, will be considered. **The rank-based ANCOVA will be a two-sided test and a two-sided p-value will be presented.**

The time to event secondary endpoints will be defined as time to first development of an event and will be estimated based on the Kaplan-Meier method, along with a 90% confidence interval. **Analyses will be based on observed data, where participants who receive rescue medication will be censored at the start of rescue medication**. A stratified log-rank test adjusting for the randomization stratification factor will be performed, to compare the combined vehicle control with each treatment group.

The following text has been deleted:
Continuous secondary endpoints involving change from BL to Week 24 will be analyzed using an analysis of covariance (ANCOVA) model with BL measurement of the variable as a covariate and treatment group and the randomization stratification factor as fixed factors. ~~The normality assumption will be visualized based on the histogram of residuals.~~

Change# 18
Section 4.10.4 Exploratory Analysis
The following changes in bold have been made:

All exploratory endpoint analyses will be performed on the Intention-to-Treat Analysis Set and will be based on observed data, **where observed data after the initiation of rescue medication will be set to missing** unless specified otherwise. **The following exploratory endpoint analyses will be repeated on the PPS:**

- **Change in MV from baseline to Week 24 for 1, 3, and 6 mm regions.**
- **Change in area of retinal non-perfusion from baseline to Week 24 seen on widefield FA.**
- **Proportion of participants with intact FAZ seen on widefield FA at Week 24.**

The following text was updated (changes in bold):
Data for eyes that were treated with anti-VEGF and/or laser will be imputed with LOCF from the time of starting treatment with anti-VEGF and/or laser. **Concomitant medications in the fellow eye with ATC4CD="SL01A" will be considered for anti-VEFG treatment in the fellow eye. Additionally, ocular concomitant procedures in the fellow eye will be considered for laser treatment. The ocular concomitant procedures will be reviewed and confirmed by the clinical**

team. The earliest date among the anti-VEGF and laser treatments will be considered the time of starting treatment for LOCF.

Associations between baseline covariates and compliance ~~will~~ **may** be assessed graphically: monthly compliance ~~will~~ **may** be presented as a series of boxplots for each baseline subgroup, by treatment group and overall, allowing for the observation of compliance trends within or between groups.

Change# 19
Section 4.10.3 Secondary Efficacy Analysis
The following changes have been made in this section:
The specific reasons for meeting rescue criteria as described in protocol section 2.14.9 will also be **summarized using number and percentage by treatment group**.

Change# 20
Section 4.11.1 Adverse Events
The following summaries were added (changes in bold):

- **Ocular TEAEs in the Study Eye**
- **Ocular TEAEs in the Fellow Eye**
- **Non-Ocular TEAEs**
- Drug-related TEAEs
- TEAEs leading to discontinuation of study
- TEAEs leading to discontinuation of study drug
- **All** TESAEs
- **Ocular TESAEs in the Study Eye**
- **Ocular TESAEs in the Fellow Eye**
- **Non-Ocular TESAEs**

Change# 21
Section 6.1 Schedule of Activities
The following changes have been made:

- IP adherence review was added to non-ocular assessments
- Footnotes were added:
  - Cornea should be assessed at the beginning before dilation.
  - Adherence to the dosing regimen should be monitored continuously and discussed with the participant at each visit. Sites are expected to intervene if there is suboptimal adherence or absence of data transfer and document the intervention in the patient dashboard.

Approval Signatures

Document Name: Statistical Analysis Plan 11 Feb 2024 OTT166-201

Document Number: VV-TMF-6342212

Parexel Version Number:

System Version Number: 2 .0

Document Approvals	
Reason for signing: Approved	Name: Kathleen Russell Role: Biostatistics Date of signature: 10-Feb-2024 02:39:39 GMT+0000
Reason for signing: Approved	Name: David Tanzer Role: Sponsor Date of signature: 12-Feb-2024 04:31:19 GMT+0000