# STATISTICAL ANALYSIS PLAN

**Unique Protocol ID:** CB8025-21730

**Official Title:** A Phase 2, Double-Blind, Randomized, Placebo-Controlled Study Followed by Open-Label Extension and Safety Follow-Up Phases to Evaluate the Activity of Seladelpar in Subjects with Nonalcoholic Steatohepatitis (NASH)

**NCT number:** NCT03551522

**Protocol Version Number:** Version 2.0

**Date:** 28-FEB-2020

# STATISTICAL ANALYSIS PLAN

# A Phase 2, Double-Blind, Randomized, Placebo-Controlled Study Followed by Open-Label Extension and Safety Follow-Up Phases to Evaluate the Activity of Seladelpar in Subjects with Nonalcoholic Steatohepatitis (NASH)

| | |
|---|---|
| **Protocol Number:** | CB8025-21730 |
| **Study Phase:** | Phase 2 |
| **Investigational Product:** | Seladelpar |
| **IND Number:** | 133253 |
| **Sponsor:** | CymaBay Therapeutics, Inc.<br>7999 Gateway Blvd, Suite 130<br>Newark, CA 94560<br>U.S.A. |
| **Protocol Versions:** | V1.1 / 16-FEB-2018<br>V1.2 / 26-MAR-2018<br>V2.0 / 13-SEP-2019<br>V3.0 / 20-DEC-2019 |
| **SAP:** | V2.0 / 28-FEB-2020 |

## SAP SIGNATURE PAGE

STUDY TITLE: **A Phase 2, Double-Blind, Randomized, Placebo-Controlled Study Followed by an Open-Label Extension and Safety Follow-Up Phases to Evaluate the Activity of Seladelpar in Subjects with Nonalcoholic Steatohepatitis (NASH)**

We, the undersigned, have reviewed and approved this statistical analysis plan.

Signature Date

[redacted]

# TABLE OF CONTENTS

## SUMMARY OF CHANGES

| Version | Date | Description of Changes |
|---|---|---|
| V1.0 | 22-APR-2019 | Original signed version. |
| V2.0 | 28-FEB-2020 | Updates to reflect the analysis performed for the Week 12 topline data cut.<br>• Modify analysis of cT1 (mITT Population instead of ITT Population)<br>• Add analysis of ALT, AST, and GGT percent change from baseline (not observed change).<br>• Add subgroup analysis of lipids (subjects on lipid lowering medication and subjects on HMG COA reductase inhibitors).<br>• Add summary of AEs occurring in the 12-week period.<br><br>Also, Protocol V2.0 and V3.0 was released on 23-SEP-2019 and 20-DEC-2019, respectively, which added an open label extension (OLE; V2.0) and safety follow up (SFU; V3.0) phases. The study was terminated on 25-NOV-2019 due to atypical histological findings in the end of treatment biopsies. The SFU was added to follow those subjects with atypical findings<br><br>As a result of the study termination, a substantial amount of analyses are removed.<br><br>Add formal analysis of alkaline phosphatase, total bilirubin, and direct bilirubin.<br><br>Summaries of lab CTCAE criteria, 12-lead ECG interpretations, and physical examinations are changed from a summary by visit to a shift table summary. |

## LIST OF TERMS AND ABBREVIATIONS

| | |
|---|---|
| AE | Adverse Event |
| ALT | Alanine Aminotransferase |
| ANCOVA | Analysis of Covariance |
| APRI | AST/Platelet Ratio Index |
| AST | Aspartate Aminotransferase |
| ATC | Anatomical Therapeutic Chemical |
| C4 | 7-α-hydroxy-4-cholesten-3-one |
| CI | Confidence Interval |
| BMI | Body Mass Index |
| DB | Double-Blind |
| ECG | Electrocardiogram |
| eCRF | Electronic Case Report Form |
| ELF | Enhanced Liver Fibrosis Score |
| ET | Early Termination |
| FAO | Fatty Acid Oxidation |
| FIB-4 | Fibrosis-4 |
| GGT | Gamma-glutamyl Transferase |
| HA | Hyaluronic Acid |
| HDL-C | High Density Lipoprotein Cholesterol |
| HOMA-IR | Homeostasis Model Assessment of Insulin Resistance |
| hs-CRP | High Sensitivity C-Reactive Protein |
| ITT | Intent-to-Treat |
| LDL-C | Low Density Lipoprotein Cholesterol |
| LLN | Lower Limit of Normal |
| mITT | Modified Intent-to-Treat |
| mITTb | Modified Intent-to-Treat Population with Biopsy |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MMRM | Mixed-Model Repeated-Measures |
| MRE | Magnetic Resonance Elastrography |
| MRI-PDFF | Magnetic Resonance Imaging-estimated Proton Density Fat Fraction |
| NAFLD | Nonalcoholic Fatty Liver Disease |
| NAS | NAFLD Activity Score |
| NASH | Nonalcoholic Steatohepatitis |
| NCI CTCAE | National Cancer Institute Common Terminology Criteria for Adverse Events |
| Non-HDL | Non High Density Lipoprotein Cholesterol |
| OLE | Open-Label Extension |
| PIIINP | Procollagen III N-Terminal Peptide |
| PP | Per-Protocol |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SFU | Safety Follow-Up |
| TC | Total Cholesterol |
| TEAE | Treatment Emergent Adverse Event |
| TIMP-1 | Tissue Inhibitor of Matrix Metaloproteinases |
| ULN | Upper Limit of Normal |
| WHO | World Health Organization |

# 1. INTRODUCTION

This document provides a description of the statistical methods and procedures to be implemented for the analyses of data from the study with protocol number CB8025-21730. Any deviations from this Statistical Analysis Plan (SAP) will be documented in the final clinical study report.

## 1.1 Study Termination

On November 25, 2019, CymaBay Therapeutics announced that the company was terminating the study. The decision to terminate that study based on atypical histological findings observed in the first 86 centrally read, end of treatment biopsies from study subjects. Planned, blinded histological assessments of the first tranche of liver biopsies in the trial revealed atypical histological findings, including histology characterized as an interface hepatitis presentation, with or without biliary injury.

At the time of study termination, 181 subjects were randomized and enrollment had been closed. Eleven subjects had been dosed in the open-label extension (OLE) phase under protocol V2.0 (13-SEP-2019). Week 12 topline primary efficacy analysis was conducted per protocol, after a database freeze of the necessary electronic case report forms and prior to study cancellation and unblinding, in June 2019.

In order to protect patient safety, and aid further investigation to better understand the findings leading to study termination, the study was unblinded December 11, 2020 prior to database lock of the double-blind (DB) and OLE phases of the study. There will be two database locks: the first for the DB and OLE phases of the study, and the second after the completion of the safety follow-up (SFU) phase. This SAP describes analyses for DB, OLE, and SFU phases, and will be finalized prior to the main study DB database lock (DB and OLE phases) and the analyses will be conducted after the respective database lock.

# 2. STUDY OBJECTIVES AND STUDY DESIGN

## 2.1 Study Objectives

### 2.1.1 Primary Objectives

1. To evaluate the effect of seladelpar on hepatic fat fraction, as assessed by magnetic resonance imaging-proton density fat fraction (MRI-PDFF) at Week 12 in the DB phase
2. To evaluate the safety and tolerability of seladelpar in subjects with nonalcoholic steatohepatitis (NASH) in the DB, OLE, and SFU phases

### 2.1.2 Secondary Objectives

1. To evaluate the effect of seladelpar on hepatic fat fraction, as assessed by MRI-PDFF at Week 26 and Week 52 in the DB phase
2. To evaluate the effect of seladelpar on histological improvement of nonalcoholic fatty liver disease (NAFLD) activity score (NAS) at Week 52 in the DB phase
3. To evaluate the effect of seladelpar on histological improvement of fibrosis at Week 52 in the DB phase
4. To evaluate the effect of seladelpar on metabolic biochemical markers and biochemical markers of inflammation over 52 weeks of treatment in the DB and OLE phases

### 2.1.3 Exploratory Objectives

1. To evaluate the effect of seladelpar on biochemical markers of fibrosis in the DB and OLE phases
2. To evaluate the effect of seladelpar on fibrosis and inflammation as assessed by magnetic resonance elastography (MRE) and LiverMultiScan in the DB phase
3. To evaluate aspartate aminotransferase (AST)/platelet ratio index (APRI) and Fibrosis-4 (FIB-4) score in the DB and OLE phases
4. To assess markers of target engagement as measured by 7-α-hydroxy-4-cholesten-3-one (C4), serum bile acids and fatty acid oxidation (FAO tests) in the DB and OLE phases
5. To evaluate histology, biochemical markers and imaging parameters during the SFU phase in subjects with significant histologic findings on the Week 52 or Early Termination (ET) liver biopsy

## 2.2 Study Design

**DB Phase**

This phase of the study is designed as a dosing ranging proof-of-concept study to evaluate the safety and potential efficacy effects of seladelpar on MRI-PDFF. The study includes liver biopsies to confirm that subjects have histological evidence of NASH at screening and to evaluate histological improvement in NASH and fibrosis after 52 weeks of treatment.

The study is planning to enroll approximately 175 male and female subjects. Subjects will be randomly assigned to receive placebo, seladelpar 10 mg/day, seladelpar 20 mg/day or seladelpar 50 mg/day (1:2:2:2 ratio). Subjects will be stratified by diabetic status (yes/no) and fibrosis stage (F1 versus F2-3). Study drug (placebo or seladelpar) will be taken in a blinded manner orally once a day for a period of 52 weeks.

The screening period will be up to 10 weeks, the blinded treatment period will be up to 52 weeks, and a follow-up period of 4 weeks after the last dose of study drug (i.e. Week 56). During the treatment period, the first on-treatment visit will occur 4 weeks after initiation of the study drug. Subjects will then be evaluated in clinic every 4 weeks until Week 12. After Week 12, visits will be 6-7 weeks apart. Subjects will also be contacted by phone or email at Weeks 19, 32 and 45. The amount of fat in the liver will be evaluated by MRI-PDFF at baseline, Weeks 12, 26, and 52. A liver biopsy will be performed twice during the DB phase: at baseline and at Week 52.

The primary efficacy analysis will be the change from baseline to Week 12 in MRI-PDFF. The study will continue in a blinded fashion until the Week 52 biopsy is collected.

The schedule of assessments for the DB phase can be found in Protocol Table 1.

**OLE Phase**

All subjects who participate in the DB phase, complete the 52-week biopsy and the 56-week follow-up labs, and meet the specified enrollment criteria will be offered participation in the OLE phase of the study. Subjects will not be required to have any additional liver histology, imaging or elastography assessments for enrollment or on-treatment assessments of treatment effect. The OLE Day 1 visit may be combined with the Week 56 visit in the DB phase. Informed Consent may occur at Week 52 or Week 56 in the DB phase for subjects still active in the DB phase or any time prior to any OLE procedures in patients who have completed the DB phase. Eligibility assessment

may occur on or after the Week 56 in the DB phase. Laboratory assessment for OLE suitability may be required if more than 45 days have elapsed since last lab collection to anticipated first OLE dosing. If greater than 45 days have elapsed since the last lab collection, subjects must complete an unscheduled visit where all labs collected at Week 56 will be retested. Subjects must also be clinically stable based on the investigator's medical judgement and key enrollment criteria to enter the OLE. All eligible subjects will receive seladelpar 50 mg regardless of dose received during the DB phase. Clinic visits during the 52-week OLE phase will occur at Day 1 and Weeks 2, 13, 26, 39 and 52 with a final EOS visit at approximately Week 56 visit. The primary objective of the OLE phase is to assess long-term safety and treatment effect as measured by biochemical tests after an additional 52 weeks of treatment.

The schedule of assessments for the OLE phase can be found in Protocol Table 2.

**SFU Phase**

All subjects who participated in the DB phase that have significant histologic findings other than NASH on the Week 52 or ET liver biopsy and received active seladelpar for at least 6 months will be asked to enroll in the SFU phase. Subjects in the placebo dosing group will not be enrolled into the SFU phase. After informed consent is obtained, an SFU visit will be performed at a minimum of 6 months after the Week 52 or ET liver biopsy. A repeat liver biopsy will be obtained to reassess the status of the original significant histologic findings. Other study procedures performed during the visit include a physical exam, vital signs, safety laboratory assessments, exploratory liver injury biomarker assessments, and FibroScan. Subjects with resolution or stabilization of the observed histologic findings will have completed the SFU phase, and no further visits or assessments will be required. Subjects worsening of the observed histologic findings will be followed up by phone at 3-month intervals and may be requested to have additional SFU visits with a repeat of some or all of the above procedures.

## 2.3 Study Outcome Measures

### 2.3.1 Efficacy Outcome Measures

The primary efficacy outcome measure is the relative change (i.e. percent change) in MRI-PDFF at Week 12 in the DB phase.

The secondary efficacy outcome measures related with MRI-PDFF are:

- observed change in MRI-PDFF at Weeks 12, 26, and 52 in the DB phase;
- relative change (i.e. percent change) in MRI-PDFF at Weeks 26 and 52 in the DB phase; and
- proportion of subjects with a relative decrease in MRI-PDFF ≥ 30% (i.e. percent change ≤ -30%) at Weeks 12, 26, and 52 in the DB phase.
- proportion of subjects with an observed MRI-PDFF change ≥ 5% (reduction, i.e. change ≤ -5%) at Weeks 12, 26, and 52 in the DB phase.

The secondary efficacy outcome measures related with liver biopsy include the following:

- proportion of subjects with reversal of NASH and no worsening of hepatic fibrosis (centrally scored histology at 52 weeks) in the DB phase. The reversal of NASH is defined as the absence of hepatocellular ballooning (score of 0) and no or minimal inflammation (lobular inflammation score of 0 or 1);
- proportion of subjects with improvement by at least one stage in fibrosis without worsening of NASH (i.e. improvement by at least one fibrosis stage without worsening of NAFLD activity score) in the DB phase; and
- proportion of subjects with a ≥ 2-point improvement in NAS in the DB phase.

Other secondary efficacy outcome measures include the following:

- relative (percent) and observed change in alanine aminotransferase (ALT), AST, gamma-glutamyl transferase (GGT) in the DB and OLE phases;
- relative (percent) and observed change in total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C), non-high-density lipoprotein cholesterol (non-HDL), homeostatic model assessment of insulin resistance (HOMA-IR) in the DB and OLE phases; and
- relative (percent) and observed change in high-sensitivity C-reactive protein (hs-CRP) in the DB and OLE phases.

The exploratory efficacy variables include the following:

- MRE changes at Week 52 in the DB phase; and
- Liver*MultiScan* (iron corrected T1 – cT1) changes at Weeks 12, 26, and 52 in the DB phase.

### 2.3.2 Safety Outcome Measures

The primary safety outcome measures include adverse events (AEs) and treatment emergent adverse events (TEAEs), physical exams, vital signs, electrocardiograms (ECGs), biochemistry, and hematology (National Cancer Institute Common Terminology Criteria for Adverse Events [NCI CTCAE], Version 5.0) and pre-defined laboratory changes of interest.

## 3. STATISTICAL METHODOLOGY

### 3.1 Baseline, Endpoint, and Other Statistical Considerations

Unless otherwise stated, only data collected during the DB phase will be summarized. Data collected during the OLE and SFU phases will be listed only.

**Baseline Definition**

Baseline for the DB phase for all efficacy and safety variables will be defined as the last scheduled value obtained prior to the first dose of study drug. If the measurement at this visit is missing, the last measurement prior to the first dose of study drug will be used as baseline. The scheduled baseline visit for MRI-PDFF, FibroScan, Liver*MultiScan*, MRE, and liver biopsy is Screening. For all other efficacy and safety variables, the scheduled baseline visit is Day 1 pre-dose.

**Summary Statistics**

Descriptive statistics (number of subjects, mean, standard deviation, median, minimum, and maximum) will be used to summarize the continuous data. Discrete measures will be summarized using counts and percentages. All data collected as part of the study database will be presented in data listings.

**Efficacy Analysis Visit Windows for the DB Phase**

For efficacy parameters with limited scheduled visits (MRI-PDFF, FibroScan, Liver*MultiScan*, MRE, and liver biopsy), analysis visits will be assigned according to the labeled visit. Note that for subjects who early term, MRI-PDFF, Liver*MultiScan*, MRE, and liver biopsy will only be obtained if the withdrawal is after Week 26. For MRI-PDFF, Liver*MultiScan*, and MRE, if the early term collection is >14 days after the last dose of study drug, it will not be used in analyses. For liver biopsy, all early term results will be used regardless of when the last dose of study drug was.

For efficacy parameters with regular scheduled visits (ALT, AST, GGT, etc.), analysis visits will be assigned according to the actual study day, calculated as the assessment date – first dose date +1. The analysis visit windows will be defined as the halfway point between the target visit days.

| Analysis Visit | Target Visit Day | Analysis Visit Window (Days) |
|---|---|---|
| Baseline | 1 | ≤1 |
| Week 4 | 28 | 2 – 42 |
| Week 8 | 56 | 43 – 70 |
| Week 12 | 84 | 71 – 133 |
| Week 26 | 182 | 134 – 227 |
| Week 39 | 273 | 228 – 318 |
| Week 52 (EOT) | 364 | 319 to ≤ last dose date +14 |
| Week 56 (EOS) | 392 | > last dose date +14 |

If there is more than one assessment within a visit window, then the analysis visit will be assigned by the following priorities:

1. Use the visit with the matching visit label
2. Use the visit closest to the target visit day. In the case of ties, use the earlier visit.

Safety summaries by visit will be done according to the visit label recorded in the database, without analysis visit windowing. However, some Week 56 (EOS) visits were recorded as unscheduled visits. Windowing will be applied to these so that it may be summarized as Week 56 (EOS).

## 3.2 Analysis Populations

The safety analysis will be conducted on the Safety Population. In the event of treatment allocation errors, subjects will be analyzed for safety according to the treatment they received. Efficacy analyses will be conducted on the Intent-to-Treat (ITT) Population, the Modified Intent-to-Treat (mITT) Population, and the Modified Intent-to-Treat Population with biopsy (mITTb). In the event of treatment allocation errors, subjects will be analyzed for efficacy according to the treatment to which they were randomized. The mITT Population will be used for the primary efficacy analysis.

All analyses except MRI-PDFF will be repeated on the following subgroups:

- Subjects in the mITTb Population with atypical biopsy findings at Week 52/ET
- Subjects in the mITTb Population without atypical biopsy findings at Week 52/ET

### 3.2.1 Safety Population

The Safety Population is defined as any subject who receives at least one dose of study drug.

### 3.2.2 Intent-to-Treat Population

The ITT Population is defined as any subject who is randomized and receives at least one dose of study drug.

### 3.2.3 Modified Intent-to-Treat Population

The mITT Population is defined as any subject who is randomized, receives at least one dose of study drug, and has Baseline and Week 12 MRI-PDFF.

### 3.2.4 Modified Intent-to-Treat Population with Biopsy

The mITTb Population is defined as any subject who is randomized, receives at least one dose of study drug, and has Baseline and Week 52/ET liver biopsies.

## 3.3 Subject Disposition

The reason given for each screen failure will be listed. Subject disposition will be provided for all randomized subjects. The number and percentage of subjects in each of the following disposition categories will be presented:

- Subjects who are randomized,
- Subjects in each analysis population including the biopsy subgroups,
- Subjects who complete Week 12,
- Subjects who complete the DB phase of the study,
- Subjects who withdraw from the DB phase of the study,
- Subject who dose in the OLE,
- Subjects who complete the OLE phase of the study,
- Subjects who withdraw from the OLE phase of the study, and
- Subjects entering the SFU phase (i.e. subjects that received seladelpar and had atypical biopsy findings at Week 52/ET).

For randomized subjects who withdraw from the study, the primary reason for the withdrawal will be listed and summarized. For any subject that withdrawal on or after November 25, 2019, the reason for withdrawal will be “study cancelled by Sponsor.”

### 3.4 Demographics and Baseline Characteristics

Demographics (age, sex, race, and ethnicity), body weight, height, body mass index (BMI), HbA1c, hs-CRP, and stratification groups (diabetic status and fibrosis stage) will be summarized by randomized treatment using descriptive statistics for the Safety Population and repeated for each mITT Population, mITTb Population, and biopsy subgroups.

Baseline MRI-PDFF, liver enzymes (ALT, AST, GGT), fibrosis stage, NAS, LiverMultiScan, and MRE, FibroScan, TC, LDL-C, HDL-C, triglycerides, HbA1c, and HOMA-IR will also be summarized for disease characteristics.

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) V20.1. The number and percentage of subjects with medical history will be summarized by preferred term within system organ class and by treatment group.

Alcohol consumption collected at screening will be listed.

### 3.5 Prior/Concomitant Medications

The use of any prior medication or concomitant medication will be listed for all subjects.

Medications will be coded using World Health Organization (WHO) Drug Dictionary September 2017G B3 version. The number and percentage of subjects taking concomitant medications will be summarized by preferred term within anatomical therapeutic chemical (ATC) term and by treatment group for the Safety Population, mITT Population, mITTb Population, and biopsy subgroups.

A concomitant medication in the DB phase is defined as any medication taken on or after the first dose of study drug in the DB phase (includes those that began prior to study drug dosing but were ongoing or ended on or after the first dose of study drug). Medications that start on or after the first dose of study drug in the OLE phase are excluded.

Concomitant procedures will be listed.

### 3.6 Study Drug Exposure and Compliance

Days of overall exposure to study drug in the DB phase will be summarized by treatment group for the Safety Population, mITT Population, mITTb Population, and biopsy subgroups using descriptive statistics. Exposure in days is defined as:

*date of last dose of study drug – date of first dose + 1*

If the date of last dose of study drug from the DB phase end of study electronic case report form (eCRF) is unknown, it will be imputed based on the following priorities if available for the subject:

1. If day is unknown but month and year are known, the last day of the month will be used
2. Day prior to Week 52/End of Treatment Visit from the “date of visit” eCRF
3. Day prior to Early Termination Visit from the “date of visit” eCRF
4. Day of study completion/early termination from the DB phase end of study eCRF
5. Last available visit date from the “date of visit” eCRF

In addition, a contingency table will be provided to display the number and percentage of subjects with exposure in the following categories:

- 1 to ≤4 weeks (1-28 days),
- >4 to ≤8 weeks (29-56 days),
- >8 to ≤12 weeks (57-84 days),
- >12 to ≤26 weeks (85-182 days),
- >26 to ≤39 weeks (183-273 days),
- >39 to ≤52 weeks (274-364 days), and
- >52 weeks (>364 days).

Percent compliance will be calculated as:

$$100 \times \frac{total\ number\ of\ tablets\ dispensed - total\ number\ of\ tablets\ returned}{days\ of\ exposure \times 2}$$

If bottles were dispensed but not returned, 0 returned tablets will be assumed in the calculation (i.e. this assumes all tablets were used). Percent compliance with the study drug during the 12-week treatment period in the DB phase and the whole 52-week treatment DB period will be summarized by treatment group for the Safety Population, mITT Population, mITTb Population, and biopsy subgroups using descriptive statistics. Additionally the number and percentage of subjects within each treatment group with overall compliance in the following categories will be provided: <80%, 80% to 120%, and >120%.

Study drug interruptions and planned dose changed will be listed but will not be factored into the compliance calculation.

## 3.7 Analysis of Efficacy

### 3.7.1 Descriptive Statistics

All data will be listed for the randomized subjects. Descriptive statistics of the baseline, change from baseline and/or relative (percent) change from baseline to each post-baseline visit will be presented by treatment group for each efficacy measurement.

### 3.7.2 Analyses of Hepatic Fat Fraction by MRI-PDFF

Summary statistics (number of subjects, mean, standard deviation, median, minimum, and maximum) for MRI-PDFF at all visits and the change and relative (percent) change from baseline will be provided for the ITT and mITT Population.

#### 3.7.2.1 Primary Efficacy Analysis

The primary efficacy outcome measure is the relative change (i.e. percent change) in MRI-PDFF at Week 12.

*Relative (percent) change in MRI-PDFF = 100 × (Week 12 – Baseline) / Baseline.*

An analysis will be performed on the relative (percent) change from baseline to Week 12 in MRI-PDFF using analysis of covariance (ANCOVA) model with treatment and stratification groups (diabetic status and fibrosis stage) as factors and baseline MRI-PDFF as a covariate using the mITT Population and repeated based on the PP Population with MRI-PDFF. The LS means, standard errors, 95% confidence intervals (CIs), and p-values will be provided. All statistical comparisons will be made at a two-sided alpha of 0.05. Adjusting for multiplicity, treatment

comparisons of the primary efficacy endpoint will be tested using a hierarchical "fixed-sequence" approach to assess clinical benefit of each of the seladelpar dose levels versus placebo, starting first with the 50 mg dose at a 5% alpha level. If positive, the seladelpar 20 mg dose will be compared to placebo at a 5% alpha level. If both seladelpar 50 mg and 20 mg doses are statistically superior to placebo then the testing will continue to the 10 mg dose vs placebo. Regardless of significance level achieved, p-values will be presented in the summary tables for informational purposes but claims of statistical significance will not be made unless the rules of the hierarchical testing approach are achieved.

The SAS sample code is listed:

```
*********************************************************************
  TREATMENT: 0 (placebo), 1 (seladelpar 10mg), 2 (seladelpar 20mg), or
             3 (seladelpar 50mg)
  DIABSTAT: diabetic status (yes or no)
  FSTAGE: Fibrosis Stage (F1 or F2-3)
  BASE: baseline MRI-PDFF
  PCHG: percent change from baseline to Week 12 in MRI-PDFF
*********************************************************************;
proc glm data=ADEFF;
  class TREATMENT DIABSTAT FSTAGE;
  model PCHG = TREATMENT DIABSTAT FSTAGE BASE;
  lsmeans TREATMENT / e pdiff cl;
  estimate "Seladelpar 10mg : Placebo" TREATMENT -1 1 0 0;
  estimate "Seladelpar 20mg : Placebo" TREATMENT -1 0 1 0;
  estimate "Seladelpar 50mg : Placebo" TREATMENT -1 0 0 1;
run;
```

The p-value from the Shapiro-Wilk normality test of the residuals from the model above will be provided. If it is <0.01, the ANCOVA model specified above will be performed on the log-transformed data, and the results will be back-transformed to relative (percent) change. If the residuals of this transformed model are still not normal, then non-parametric analyses will be explored.

#### 3.7.2.2 Hepatic Fat Fraction Treatment Goals

The secondary efficacy outcome measures related to MRI-PDFF treatment goals are:

- proportion of subjects with a relative decrease in MRI-PDFF ≥ 30% (i.e. percent change ≤ -30%) at Weeks 12, 26, and 52;
- proportion of subjects with an observed MRI-PDFF change ≥ 5% (reduction, i.e. change ≤ -5%) at Weeks 12, 26, and 52.

For each endpoint, the number and percentage of responders will be provided by treatment group and by stratification group. Cochran-Mantel-Haenszel tests will be performed adjusting for stratification groups (diabetic status and fibrosis stage). The general association p-value will be provided. Then for each seladelpar treatment group compared to placebo, the odds ratio, 95% confidence interval, and Breslow-Day test for homogeneity of the odds ratio p-value will be provided. The analysis will be performed on the ITT Population and repeated for the mITT Population. Subjects with missing data will be considered unimproved on the outcome measure.

```
*********************************************************************
  TREATMENT: 0 (placebo), 1 (seladelpar 10mg), 2 (seladelpar 20mg), or
```

```
            3 (seladelpar 50mg)
  DIABSTAT: diabetic status (yes or no)
  FSTAGE: Fibrosis Stage (F1 or F2-3)
  RESPONDER: yes or no
*********************************************************************;
proc freq data=ALT;
   tables DIABSTAT*FSTAGE*TREATMENT*RESPONDER / cmh;
   weight COUNT;
run;
```

### 3.7.3 Analyses of Liver Biopsy Results

All endpoints will be summarized and analyzed based on the mITTb Population and biopsy subgroups.

Responder outcome measures include:

- proportion of subjects with reversal of NASH and no worsening of hepatic fibrosis (centrally scored histology at 52 weeks). The reversal of NASH is defined as the absence of hepatocellular ballooning (score of 0) and no or minimal inflammation (lobular inflammation score of 0 or 1);
- proportion of subjects with improvement by at least one stage in fibrosis without worsening of NASH (i.e. improvement by at least one fibrosis stage without worsening of NAFLD activity score); and
- proportion of subjects with a ≥ 2-point improvement in NAS.

For each endpoint, the number and percentage of responders will be provided by treatment group and by stratification group. Cochran-Mantel-Haenszel tests will be performed adjusting for stratification groups (diabetic status and fibrosis stage). The general association p-value will be provided. Then for each seladelpar treatment group compared to placebo, the odds ratio, 95% confidence interval, and Breslow-Day test for homogeneity of the odds ratio p-value will be provided. Subjects with missing data will be considered unimproved on the outcome measure.

### 3.7.4 Analyses of Other Efficacy Endpoints

All endpoints will be summarized and analyzed based on the ITT Population and repeated on the mITTb Population and biopsy subgroups. Since these endpoints are for exploratory purposes, no adjustments for multiplicity will be made. Summaries will exclude OLE phase visits. All OLE data will be listed.

#### 3.7.4.1 ALT, AST, and GGT

The relative (percent) change from baseline in ALT, AST, and GGT at Weeks 4, 8, 12, 26, 39, 52, and 56 will be analyzed using a MMRM. No imputation will be performed. The factors in the model will be stratification groups (diabetic status and fibrosis stage), treatment group, baseline value, visit, and the treatment group by visit interactions. An unstructured covariance matrix will be used (TYPE=UN).

#### 3.7.4.2 TC, HDL-C, LDL-C, Non-HDL, and HOMA-IR

The percent change from baseline in TC, HDL-C, LDL-C, Non-HDL, and HOMA-IR at Weeks 4, 8, 12, 26, 39, 52, and 56 will be analyzed using a MMRM. No imputation will be performed. The factors in the model will be stratification groups (diabetic status and fibrosis stage), treatment group, visit, and the treatment group by visit interactions. An unstructured covariance matrix will be used (TYPE=UN). The same model will be repeated on the subgroups of subjects on lipid lowering medication and HMG COA reductase inhibitors.

#### 3.7.4.3 Triglycerides

The relative (percent) change from baseline in triglycerides at Weeks 4, 8, 12, 26, 39, 52, and 56 will be analyzed using a MMRM on the log-transformed data. No imputation will be performed. The factors in the model will be stratification groups (diabetic status and fibrosis stage), treatment group, baseline value, visit, and the treatment group by visit interactions. An unstructured covariance matrix will be used (TYPE=UN). The same model will be repeated on the subgroups of subjects on lipid lowering medication and HMG COA reductase inhibitors.

#### 3.7.4.4 Hs-CRP

The relative (percent) change from baseline in hs-CRP at Weeks 4, 8, 12, 26, 39, 52, and 56 will be analyzed using a MMRM on the log-transformed data. No imputation will be performed. The factors in the model will be stratification groups (diabetic status and fibrosis stage), treatment group, baseline value, visit, and the treatment group by visit interactions. An unstructured covariance matrix will be used (TYPE=UN).

#### 3.7.4.5 Other Efficacy Variables

Liver*MultiScan*

The observed change and relative (percent) change from baseline in Liver*MultiScan* (cT1 whole median) will be summarized.

Fibroscan

Results from the FibroScan, if available, will be summarized.

## 3.8 Analysis of Safety

The safety endpoints for this study include: AEs, safety laboratory assessments, physical examinations, vital signs, and ECGs. The safety endpoints will be summarized based on the Safety Population and repeated on the mITTb Population and biopsy subgroups. Missing values will not be imputed. Safety data will be summarized by study treatment (placebo vs. seladelpar) and by seladelpar dose.

### 3.8.1 Adverse Events

An AE is any medical occurrence in a subject administered a pharmaceutical product in a clinical study, regardless of a causal relationship with this treatment. An AE can, therefore, be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease

temporally associated with the use of a medicinal product, whether or not related to the medicinal product.

For analysis, a TEAE is defined as any AE that occurred after the first dose of double-blind study drug and up to 30 days after last dose of double-blind study drug, or one day before the first dose of OLE phase, whichever is earlier.

Only AEs in the DB phase will be included in the summary tables. AEs beginning after the first dose of the OLE phase will be excluded from summaries but will be included in listings.

A summary overview of TEAEs will be provided, which presents the number and percentage of subjects in each treatment group from the Safety Population, including number of events, satisfying each of the following categories:

- Any TEAEs,
- Maximum severity of TEAEs,
- Study drug-related TEAEs,
- Maximum severity of study drug-related TEAEs,
- All serious AEs (SAEs),
- All study drug-related SAEs,
- AEs leading to death,
- AEs leading to dose decreased or temporary interruption,
- AEs leading to study drug discontinuation, and
- Study drug-related AEs leading to study drug discontinuation.

Adverse events will be coded using MedDRA. The number and percentage of subjects with TEAEs, including the number of TEAEs, will be summarized by their MedDRA preferred term within system organ class and by treatment. For subject count summaries, multiple AE events with the same MedDRA coded terms (preferred term and system organ class) from the same subject will only be counted once. Similar summaries will be provided for drug-related TEAEs.

The number and percentage of subjects with TEAEs will be summarized by reported maximum severity within each MedDRA preferred term within system organ class and by treatment. Similar summaries will be provided for SAEs, for drug-related TEAEs, and for TEAEs beginning in the DB phase 12 week treatment period.

Listings will be provided of all AEs, including SAEs and AEs leading to study drug discontinuation.

### 3.8.2 Clinical Laboratory Assessments

#### 3.8.2.1 Chemistry, Hematology, and Other lab assessments

Descriptive statistics of each chemistry (including HbA1c) and hematology parameter will be presented for baseline values and for values and the relative (percent) changes and changes from baseline at each post-baseline visit. These will be presented by treatment group for each parameter using the Safety Population.

The relative (percent) change from baseline in alkaline phosphatase, total bilirubin, and direct bilirubin at each scheduled treatment visit will be analyzed using a MMRM. No imputation will be performed. The factors in the model will be stratification groups (diabetic status and fibrosis

stage), treatment group, baseline value, visit, and the treatment group by visit interactions. An unstructured covariance matrix will be used (TYPE=UN).

Counts and percentages of subjects with meeting the following categories at any time post-dose in the DB phase will be tabulated.

- Elevated ALT
- Elevated AST
- Elevated Total Bilirubin
- CK >2.5xULN
- Lipase >3xULN
- Toponin I >1xULN
- Creatinine >1.25xULN
- INR >1.4
- Platelets <1xLLN (or decreased platelet estimate)

Subjects who met the safety monitoring criteria (protocol section 8.4) will be listed.

Counts and percentages of subjects with chemistry parameters meeting the following categories based on NCI CTCAE will be tabulated as a shift table from baseline to the worst post-dose result in the DB phase. Results after dosing in the OLE phase will be excluded.

| Criteria | Grade 1 | Grade 2 | Grade 3 | Grade 4 |
|---|---|---|---|---|
| Alanine aminotransferase increased | >ULN - 3.0 x ULN if baseline was normal (i.e. if baseline was ≤1.0 x ULN);<br>1.5 - 3.0 x baseline if baseline was abnormal (i.e. if baseline was >1.0 x ULN) | >3.0 - 5.0 x ULN if baseline was normal;<br>>3.0 - 5.0 x baseline if baseline was abnormal | >5.0 - 20.0 x ULN if baseline was normal;<br>>5.0 - 20.0 x baseline if baseline was abnormal | >20.0 x ULN if baseline was normal;<br>>20.0 x baseline if baseline was abnormal |
| Aspartate aminotransferase increased | >ULN - 3.0 x ULN if baseline was normal;<br>1.5 - 3.0 x baseline if baseline was abnormal | >3.0 - 5.0 x ULN if baseline was normal;<br>>3.0 - 5.0 x baseline if baseline was abnormal | >5.0 - 20.0 x ULN if baseline was normal;<br>>5.0 - 20.0 x baseline if baseline was abnormal | >20.0 x ULN if baseline was normal;<br>>20.0 x baseline if baseline was abnormal |
| Blood bilirubin increased | Total Bilirubin<br>>ULN - 1.5 x ULN if baseline was normal;<br>> 1.0 - 1.5 x baseline if baseline was abnormal | Total Bilirubin<br>>1.5 - 3.0 x ULN if baseline was normal;<br>>1.5 - 3.0 x baseline if baseline was abnormal | Total Bilirubin<br>>3.0 - 10.0 x ULN if baseline was normal;<br>>3.0 - 10.0 x baseline if baseline was abnormal | Total Bilirubin<br>>10.0 x ULN if baseline was normal;<br>>10.0 x baseline if baseline was abnormal |
| Creatinine increased | >ULN - 1.5 x ULN | >1.5 - 3.0 x baseline<br>or<br>>1.5 - 3.0 x ULN | >3.0 x baseline or<br>>3.0 - 6.0 x ULN | >6.0 x ULN |

| Criteria | Grade 1 | Grade 2 | Grade 3 | Grade 4 |
|---|---|---|---|---|
| CPK increased | >ULN - 2.5 x ULN | >2.5 x ULN - 5 x ULN | >5 x ULN - 10 x ULN | >10 x ULN |
| Hypoalbuminemia | <LLN - 3 g/dL;<br><LLN - 30 g/L | <3 - 2 g/dL;<br><30 - 20 g/L | <2 g/dL;<br><20 g/L | |
| Hyponatremia (Low Sodium) | <LLN - 130 mmol/L | 125-129 mmol/L | 120-124 mmol/L | <120 mmol/L |
| Hypernatremia (High Sodium) | >ULN - 150 mmol/L | >150 - 155 mmol/L | >155 - 160 mmol/L | >160 mmol/L |
| Serum amylase increased | >ULN - 1.5 x ULN | >1.5 - 2.0 x ULN | >2.0 - 5.0 x ULN | >5.0 x ULN |
| Lipase increased | >ULN - 1.5 x ULN | >1.5 - 2.0 x ULN; | >2.0 - 5.0 x ULN | >5.0 x ULN |

Counts and percentages of subjects with hematology parameters meeting the following categories based on NCI CTCAE will be tabulated as a shift table from baseline to the worst post-dose result in the DB phase. Results after dosing in the OLE phase will be excluded.

| Criteria | Grade 1 | Grade 2 | Grade 3 | Grade 4 |
|---|---|---|---|---|
| Anemia | Hemoglobin<br><LLN - 10.0 g/dL;<br><LLN - 6.2 mmol/L;<br><LLN - 100 g/L | <10.0 - 8.0 g/dL;<br><6.2 - 4.9 mmol/L;<br><100 – 80 g/L | <8.0 g/dL;<br><4.9 mmol/L;<br><80 g/L | |
| Hemoglobin increased | Increase in >0 - 2 g/dL above ULN or above baseline if baseline is above ULN | Increase in >2 - 4 g/dL above ULN or above baseline if baseline is above ULN | Increase in >4 g/dL above ULN or above baseline if baseline is above ULN | |
| INR increased | >1.2 - 1.5;<br>>1 - 1.5 x baseline if on anticoagulation | >1.5 - 2.5;<br>>1.5 - 2.5 x baseline if on anticoagulation | >2.5;<br>>2.5 x baseline if on anticoagulation | |
| Neutrophil count decreased | <2000 - 1500/mm3;<br><2.0 - 1.5 x 10e9 /L | <1500 - 1000/mm3;<br><1.5 - 1.0 x 10e9 /L | <1000 - 500/mm3;<br><1.0 - 0.5 x 10e9 /L | <500/mm3;<br><0.5 x 10e9 /L |
| White blood cells decreased | <LLN - 3000/mm3;<br><LLN - 3.0 x 10e9 /L | <3000 - 2000/mm3;<br><3.0 - 2.0 x 10e9 /L | <2000 - 1000/mm3;<br><2.0 - 1.0 x 10e9 /L | <1000/mm3;<br><1.0 x 10e9 /L |
| Platelet count decreased | <LLN - 75,000/mm3;<br><LLN - 75.0 x 10e9 /L | <75,000 - 50,000/mm3;<br><75.0 - 50.0 x 10e9 /L | <50,000 - 25,000/mm3;<br><50.0 - 25.0 x 10e9 /L | <25,000/mm3;<br><25.0 x 10e9 /L |

### 3.8.3 12-Lead Electrocardiogram

Overall ECG interpretations (normal, abnormal not clinically significant, abnormal clinically significant) will be tabulated as a shift table from baseline to the worst post-dose result in the DB phase. Results after dosing in the OLE phase will be excluded.

### 3.8.4 Physical Examination

Physical examination findings will be tabulated as a shift table from baseline to the worst post-dose result in the DB phase. Results after dosing in the OLE phase will be excluded

### 3.8.5 Vital Signs, Weight, BMI

Vital signs parameters (systolic blood pressure, diastolic blood pressure, heart rate, temperature, respiratory rate), weight, and BMI will be summarized using descriptive statistics for the Safety Population. The relative (percent) change from baseline will also be presented.

### 3.8.6 Other Safety Parameters

Other safety data will be listed.

## 3.9 Safety Follow-up Phase

After the final database lock, the study data listings will be updated to include all data that is collected during the SFU phase. Summary of baseline characteristics, demographics and safety data of subjects who participate SFU phase may be summarized if applicable.

# 4. SAMPLE SIZE DETERMINATION

A sample size of 175 subjects (50 subjects per each dose of seladelpar and 25 subjects for placebo) provides at least 80% power to detect a pairwise treatment difference of at least 20% between the active and placebo groups with respect to a relative change from baseline in hepatic fat fraction, as assessed by MRI-PDFF at Week 12. Power calculations were performed using a standard t-test, allowed for a 12% dropout rate (6 subjects per seladelpar arm, 3 subjects for the placebo arm), and a pooled standard deviation of 26 for the relative change endpoint.

# 5. CHANGES FROM PROTOCOL

The Protocol specified use of NCI CTCAE version 4.0. However, version 5.0 has become available and will be used for summaries.

Due to the early termination of the study, the following will not be analyzed:

- MMRM and Wilcoxon rank-sum analysis of MRI-PDFF
- proportion of subjects with normalization of MRI-PDFF (defined as a fat fraction of < 5%) at Weeks 12, 26, and 52 in the DB phase
- CMH analysis of biopsy endpoints
- Summaries of APRI and FIB-4 in the DB and OLE phases (all reported data will still be listed)
- Summaries of fibrinogen in the DB and OLE phases (all reported data will still be listed)
- Summaries of enhanced liver fibrosis score (ELF) and its components (hyaluronic acid [HA], procollagen III N-terminal peptide [PIIINP], tissue inhibitor of matrix metalloproteinases [TIMP-1]) in the DB and OLE phases (all reported data will still be listed)
- Summaries of C4, serum bile acids, and FAO tests in DB and OLE phases (all reported data will still be listed)

- PP Population

## 6. PROGRAMMING SPECIFICATIONS

Statistical analyses will be performed using SAS® version 9.4. The Medpace standard operating procedures will be followed for the validation of all SAS programs and outputs. The mock-ups for SAS-generated tables/figures/listings will be prepared in a separate document and finalized before database lock for the study.