STATISTICAL ANALYSIS PLAN

Study Title:	A Proof-of-Concept and Dose-Ranging Study Investigating the Efficacy and Safety of HTD1801 in Adults with Nonalcoholic Steatohepatitis (NASH) and Type 2 Diabetes Mellitus (T2DM)
Phase:	2
Protocol No.:	HTD1801.PCT012
Protocol Date(s):	Protocol Version 3.0, 27MAR2019
Analysis Plan Version and Date:	Final Version 4, 25MAR2020
Prepared By:	Alison Miller PharPoint Research, Inc. 5003 S. Miami Bvd. Suite 100 Durham, NC 27703
Prepared For:	HighTide Biopharma Pty. Ltd. 1 Melissa Street Mount Waverley, Vic 3149 Australia

CONFIDENTIAL AND PROPRIETARY INFORMATION

STATISTICAL ANALYSIS PLAN REVIEW AND APPROVAL

This Statistical Analysis Plan has been prepared in accordance with team reviewers' specifications.

Prepared by:

DocuSigned by:
Alison Miller
Signer Name: Alison Miller
Signing Reason: I am the author of this document
Signing Time: 25-Mar-2020 | 17:07 PM EDT
A45D42F79F1E4B78942F5F202E6B2B75

25-Mar-2020 | 17:07 PM EDT

Alison Miller, MStat
Biostatistician
PharPoint Research, Inc.

Date

Review:

DocuSigned by:
Stephanie Moody
Signer Name: Stephanie Moody
Signing Reason: I approve this document
Signing Time: 25-Mar-2020 | 17:27 PM EDT
A9C789155BCC45BDB13AF80CFDB8C2E2

25-Mar-2020 | 17:27 PM EDT

Stephanie Moody, MStat
Senior Manager, Biostatistics
PharPoint Research, Inc.

Date

Cathryn Bennett

25 Mar 20 2:00pm PACIFIC

On behalf of HighTide Biopharma Pty. Ltd. (HighTide):
Cathryn Bennett, B.N., R.N., C.C.R.A.
VP, Clinical Operations
HighTide

Date

TABLE OF CONTENTS

1. INTRODUCTION

This document describes the statistical methods and data presentations to be used in the summary and analysis of data from Protocol HTD1801.PCT012. Background information is provided for the overall study design and objectives. The reader is referred to the study protocol and case report forms (CRFs) for details of study conduct and data collection.

1.1. STUDY OVERVIEW

Protocol HTD1801.PCT012 is a Phase 2, multicenter, randomized, double-blind, placebo-controlled, parallel-group, proof-of-concept, and dose-ranging study comparing two doses of investigational product (HTD1801) to placebo. Approximately 117 patients with Nonalcoholic Steatohepatitis (NASH) and Type 2 Diabetes Mellitus (T2DM) will be enrolled as subjects from approximately 15 investigational sites in the United States (US). Upon confirmation of eligibility, subjects will be randomized in a 1:1:1 ratio to one of the following treatment arms:

- **Treatment Arm 1**: HTD1801 500 mg twice daily (BID)
- **Treatment Arm 2**: HTD1801 1000 mg BID
- **Treatment Arm 3**: Placebo BID

Subjects will receive the treatment to which they are randomly assigned for the duration of the 18-week double-blind treatment period. Clinic visits for efficacy assessments and safety monitoring will occur at Week 2, then every 4 weeks throughout the double-blind treatment period. A follow-up safety telephone contact for an assessment of adverse events (AEs) and concomitant medications will take place at least 30 days after the last dose of study drug. An interim analysis will be conducted once a minimum of 51 subjects complete the assessment of the primary efficacy endpoint. An independent Data and Safety Monitoring Board (DSMB) will provide additional safety oversight.

1.2. SCHEDULE OF EVENTS

	SCR	Double-Blind Treatment						ET	Follow-up
		Baseline	WK 2	WK 6	WK 10	WK 14	WK 18		
Visit Day	**-28 to -1**	**0**	**14**	**42**	**70**	**98**	**126**	**-**	**+30d**
Procedure / Window	**-**	**-**	**±3d**	**±3d**	**±3d**	**±3d**	**±3d**		**+7d**
Administrative Procedures									
Clinic visit	X	X	X	X	X	X	X	X	
Informed consent	X								
Eligibility criteria	X	X[a]							
Demographics	X								
Medical history	X	X[a]							
Randomization		X							
Dispense study drug		X	X	X	X	X			
Return study drug			X	X	X	X	X	X	
Study discharge							X	X	
Telephone contact									X
Clinical Assessments									
Comprehensive physical examination	X						X	X	
Targeted physical examination[b]		X[a]		X		X			
Vital signs	X	X[a]		X		X	X	X	
Weight	X	X[a]	X	X	X	X	X	X	

	SCR	Double-Blind Treatment						ET	Follow-up
		Baseline	WK 2	WK 6	WK 10	WK 14	WK 18		
Visit Day	**-28 to -1**	**0**	**14**	**42**	**70**	**98**	**126**	**-**	**+30d**
Procedure / Window	**-**	**-**	**±3d**	**±3d**	**±3d**	**±3d**	**±3d**		**+7d**
Height	X								
ECG	X	X[a]		X		X	X	X	
Laboratory Assessments									
Serum chemistry panel[c]	X	X	X	X	X	X	X	X	
Lipid panel[c]	X	X			X		X	X	
Hematology with differential[c]	X	X			X		X	X	
Coagulation panel[c]	X	X			X		X	X	
HbA1c	X[d]				X		X	X	
HOMA-IR		X					X	X	
HIV and Hepatitis B/C	X								
G6PD	X								
Pregnancy test[e]	X	X[a,e]					X	X	
FSH test[e]	X								
Urine drug test	X								
Imaging and Biomarker Assessments									
MRI-PDFF	X[d]						X	X	
Corrected T1 and T2*	X[d]								
Pro-C3		X					X	X	

	SCR	Double-Blind Treatment						ET	Follow-up
		Baseline	WK 2	WK 6	WK 10	WK 14	WK 18		
Visit Day	**-28 to -1**	**0**	**14**	**42**	**70**	**98**	**126**	**-**	**+30d**
Procedure / Window	**-**	**-**	**±3d**	**±3d**	**±3d**	**±3d**	**±3d**		**+7d**
ELF test		X					X	X	
Bile acid panel[c]		X			X		X	X	
Safety Assessments									
Adverse event monitoring	X	X[a]	X	X	X	X	X	X	X
Prior and concomitant medications	X	X[a]	X	X	X	X	X	X	X

[a] Continued study eligibility confirmed based on updated medical history, targeted physical examination, vital signs, weight, ECG, negative urine pregnancy test (for FOCP), updated prior and concomitant medications, and any adverse events that have occurred since signing informed consent.

[b] Targeted physical examination of body systems or organs performed on symptom-driven basis (e.g., adverse events or other findings).

[c] See Table 2 of the protocol for analytes tested as part of the serum chemistry, lipid, hematology, and coagulation panels. The serum chemistry panel performed at the Week 2, Week 6, and Week 14 visits only includes testing of ALP, ALT, AST, GGT, total bilirubin, direct bilirubin, and fasting glucose. The Week 10 and Week 18 (or ET) bile acid panel applies only to subjects who have had the Baseline bile acid panel performed.

[d] Assessments performed during Screening to confirm study eligibility; results applied as Baseline values.

[e] Pregnancy test required for FOCP only; serum pregnancy test performed at Screening and Week 18 (or ET); urine and confirmatory serum pregnancy tests performed at Baseline. FSH test required for postmenopausal females only.

ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; ECG = electrocardiogram; ELF = enhanced liver fibrosis; ET = Early Termination; FOCP = females of child-bearing potential; FSH = follicle-stimulating hormone; G6PD = glucose-6-phosphate dehydrogenase; GGT = gamma-glutamyl transferase; HbA1c = hemoglobin A1c; HIV = human immunodeficiency virus; HOMA-IR = homoeostasis model assessment-estimated insulin resistance; MRI-PDFF = magnetic resonance imaging-estimated proton density fat fraction; SCR = Screening

1.3. GLOSSARY OF ABBREVIATIONS

AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BID	twice daily
BMI	body mass index
CFB	change from baseline
CRF	case report form
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	data and safety monitoring board
ECG	electrocardiogram
ELF	enhanced liver fibrosis
ET	early termination
FSH	follicle-stimulating hormone
G6PD	glucose-6-phosphate dehydrogenase
GGT	gamma-glutamyl transferase
HbA1c	hemoglobin A1c
HOMA-IR	homeostasis model assessment – estimated insulin resistance
IWRS	interactive web response system
LFC	liver fat content
MedDRA	Medical Dictionary for Regulatory Activities
MRI-PDFF	magnetic resonance imaging – estimated proton density fat fraction
NASH	nonalcoholic steatohepatitis
PP	per protocol
PT	preferred term
SAE	serious adverse event
SD	standard deviation
SOC	system organ classification
T2DM	type 2 diabetes mellitus
TEAE	treatment emergent adverse events
WHO	World Health Organization

2. OBJECTIVES

Primary Objective:

The objective of this study is to evaluate the effects of HTD1801 on liver fat content (LFC) in adults with NASH and T2DM.

Secondary Objective:

The secondary objectives are to evaluate the following:

- The effects of HTD1801 on hemoglobin A1c (HbA1c) and glucose metabolism
- The effects of HTD1801 on lipid profile
- The effects of HTD1801 on liver-associated enzymes
- The effects of HTD1801 on biomarkers of liver inflammation and fibrosis
- The effects of HTD1801 on bile acid homeostasis
- The safety and tolerability of HTD1801

3. GENERAL STATISTICAL CONSIDERATIONS

3.1. SAMPLE SIZE AND POWER

An absolute change in LFC of 5% as measured by magnetic resonance imaging-estimated proton density fat fraction (MRI-PDFF) has been reported to be the minimal clinically significant difference between active treatment and placebo that correlates with histological improvement in NASH (Loomba et al., 2015, Patel et al., 2015). Furthermore, Harrison and colleagues reported a pooled standard deviation of 6.3% for changes from Baseline in LFC (Harrison et al., 2018). Based on this standard deviation, 35 subjects in each treatment group will provide 90% power to show a difference of 5 percentage points between any 2 treatment groups at the 5% level of significance. To allow for a dropout rate of 10%, 39 subjects will be randomized to each of the 3 treatment groups. Due to the exploratory nature of this study, enrollment may be stopped early when sufficient information has been gathered to evaluate the effects of HTD1801 on LFC.

An interim analysis will be conducted after a minimum of 51 subjects complete the assessment of the primary efficacy endpoint to assess sample size assumptions and futility through conditional power.

3.2. RANDOMIZATION AND MASKING

This is a multicenter, randomized, double-blind study. Subjects will be centrally randomized via an interactive web response system (IWRS) in a 1:1:1 ratio to one of 3 groups:

HTD1801 500 mg BID

HTD1801 1000 mg BID

Placebo BID

In addition to the subject ID number, subjects will be assigned a randomization number, as well as a kit ID number corresponding to a pre-packaged and numbered study drug kit. The randomization schedule was prepared by a non-study statistician using PharPoint standard operating procedure BIO002.

All subjects and study personnel, except for the unblinded statistician, the IWRS vendor and the study drug packaging supplier, will remain blinded to the treatment administered throughout the study. All study drugs will be provided in matching white tablets, 4 tablets to a pouch, 2 pouches to be administered each day.

3.3. HANDLING OF DATA

3.3.1. Strata and Covariates

Selected efficacy analyses will include Baseline LFC and Baseline ALT as covariates.

3.3.2. Examination of Subject Subsets

Selected efficacy analyses will be presented by Baseline HbA1c group (< 7% vs. ≥ 7%).

3.3.3. Multiple Testing and Comparisons

All analyses will be conducted without adjustments for multiple comparisons.

3.3.4. Missing Data and Outliers

Every effort will be made to obtain required data at each scheduled evaluation from all subjects who have been treated. For subjects who do not complete the 18-week evaluation, endpoints that will only be measured at Week 18 or the Early Termination (ET) visit (i.e., LFC, HOMA-IR, Pro-C3, ELF) will have the value at the ET visit carried forward for the missing Week 18 visit.

For all non-LFC secondary endpoints measured at at least one post-baseline timepoint other than Week 18/ET, missing Week 18 values will be imputed using a multiple imputation procedure for all subjects with an available measurement prior to Week 18. Multiple imputation relies on the assumption that the data are missing at random. This method of imputation produces unbiased estimators of the mean and standard error, and, as such, is preferred over single imputation methods. For this method, the procedure requires production of multiple datasets containing plausible values for each missing value. These plausible values are generated using available auxiliary information (site, treatment, baseline assessment value, baseline LFC, and baseline ALT).

Using a set of values rather than a single value better accounts for the uncertainty about the value being imputed. For this imputation, the Markov chain method will be used. SAS PROC MI will be used to generate 10 possible imputed datasets.

For each variable that has imputations performed, the following steps will be followed:

1. Missing values will be imputed using PROC MI. This will result in 10 datasets per variable.
2. Efficacy endpoints will be derived from the appropriate set of 10 datasets produced by PROC MI using the imputed values.
3. Each efficacy endpoint will be analysed 10 times (once per source dataset) using the appropriate statistical model for that endpoint.
4. For each endpoint, the results from the analysis from each of the 10 datasets will be combined using PROC MIANALYZE to produce an inferential result for the parameter of interest.

3.3.5. Imputation of Incomplete Dates

An incomplete date is any date for which either the day, month, or year is unknown, but all three fields are not unknown. An incomplete date occurs when the exact date an event occurred or ended cannot be obtained from a subject. For many of the analyses, a complete date is necessary in order to determine if the event should be included in the analysis (i.e., if the event is treatment-emergent) or to establish the duration of an event. In such cases, incomplete dates will be imputed.

For purposes of imputation, all events with an incomplete end date are assumed to have ended on or before the day the form was completed. In an effort to minimize bias, the project statistician will impute dates in a systematic but reasonable manner, as decribed below.

For event start dates, if the month/year is the same as the Day 1 month/year then the date will be set to the date of Day 1. In other cases, missing days will be imputed as the day component of Day 1; missing months/years will be imputed as the month/year of Day 1.

For event end dates, if the month/year is the same as the event start date month/year then the date will be set to the date of the event start date. In other cases, missing days will be imputed as the day component of the event start date; missing months/years will be imputed as the month/year of the event start date.

3.3.6. Presentations by Study Visit

Nominal Study Visits as obtained from the CRF or applicable laboratory results will be utilized for summary displays. If assessments are collected multiple times within a given Study Visit, the result closest to the scheduled visit date will be used for summary presentations. If two measurements have the same distance to the expected date, the earlier value will be used. If a scheduled assessment and an unscheduled or ET assessment are collected within a given visit, the value from the scheduled assessment will be chosen over the value from the unscheduled assessment. ET visits will be windowed to the Week 18 visit. Unscheduled visits will be windowed based on study day, according to the table below. All assessments will be presented in the listings.

Study Day	Windowed Study Visit
11- 17	Week 2
39 - 45	Week 6
67 - 73	Week 10
95 - 101	Week 14
123 - 129	Week 18

3.3.7. Definitions and Terminology

Age
The age of a subject is defined as the number of whole years between the subject's birth date and the treatment start date.

Day 0 (Baseline)
Day 0 is the earliest day that study drug is initiated.

Study Day
Study Day is defined relative to Baseline (Day 0). Thus, the study day of an event is calculated as:

Study Day = event date – date of Day 0.

Study Visit
Study Visit is the nominal visit as recorded on the CRF.

Days on Treatment
Days on Treatment is calculated as Final Day on Treatment based on the End of Study eCRF page – First Day on Study Treatment + 1.

Baseline Value
For purposes of analysis, the baseline value is defined as the last non-missing value obtained prior to initiation of study drug.

Change from Baseline
Change from baseline for a given endpoint is defined as the Study Day X value minus the Baseline Value.

Percent (Relative) Change from Baseline
Percent, or relative, change from Baseline for a given endpoint is defined as the change from Baseline at Study Day X divided by the Baseline value, multiplied by 100 ((CFB/BASE) x 100).

Adverse Event
An AE is defined as any untoward medical occurrence associated with the use of a drug in humans, whether or not considered drug related. An AE can be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of the investigational product, regardless of whether it is considered to be related to the investigational product. All AEs will be recorded on the Adverse Event CRF.

Treatment-emergent Adverse Event
A TEAE will be an AE that occurred during the study after the first dose of study drug or that was present prior to dosing and exacerbates after the first dose of study drug. Additionally, it is assumed that an Adverse Event which was reported to have started on Day 1 without an associated onset time occurred after the initiation of study drug.

Concomitant Medications
Concomitant medications are those medications taken on or after the initiation of study drug. This definition includes medications started prior to the initiation of study drug, but continuing concomitantly with study drug.

Prior Medications
Prior medications are those medications taken prior to the initiation of study drug.

Total Bile Acid
Total bile acid is calculated as the sum of all bile acid measurements and metabolites at a visit for a given subject.

Total Primary Bile Acid
Total primary bile acid is calculated as the sum of cholic acid, chenodeoxycholic acid, glycocholic acid, glycochenodeoxycholic acid, taurocholic acid, and taurochenodeoxycholic acid at a visit for a given subject.

Total Secondary Bile Acid
Total secondary bile acid is calculated as the sum of deoxycholic acid, lithocholic acid, glycodeoxycholic acid, glycolithocholic acid, taurodeoxycholic acid, and taurolithocholic acid at a visit for a given subject.

Total Ursodeoxycholic Acid
Total ursodeoxycholic acid is calculated as the sum of ursodeoxycholic acid, glycoursodeoxycholic acid, and taurolithocholic acid at a visit for a given subject.

3.4. TIMING OF ANALYSES

An interim analysis will be conducted once source document verification is completed for a minimum of 51 subjects who complete the assessment of the primary efficacy endpoint to assess sample size assumptions and futility through conditional power. The study will not be stopped for

efficacy. The method of Cui and colleagues (Cui et al., 1999) will be used to compute the alpha spend associated with the interim analysis to assess futility and to recompute sample size by evaluating the pooled standard deviation for the LFC change from Baseline.

Conditional power calculations will be performed for the difference between each of the two active treatment groups versus placebo for the primary endpoint. The following rules will be applied for the primary efficacy endpoint:

1. If both conditional powers at the time of the interim analysis are <10%, then the study will be terminated for futility.
2. If either conditional power is >10% but both are <36%, then the sample size will not be increased and the study will continue based on the original sample size.
3. If either conditional power is >36% but both are <80%, then the sample size will be adjusted to retain the original power of 90% or a 50% increase in the sample size, whichever is smaller.
4. If either conditional power is >80%, then the study will continue as is.

The sponsor will be advised of either futility, no change, or of a specific sample size increase, but not about the degree of efficacy per treatment group. The analysis will be conducted by an independent statistician who is not otherwise involved in the operational aspects of this study.

The final analysis will be completed after the last subject completes or discontinues the study and the resulting clinical database has been cleaned, quality checked, and locked.

Additionally, safety oversight will be conducted throughout the study under the direction of a DSMB. The timing and frequency for the DSMB will be detailed in a separate DSMB charter.

4. ANALYSIS SETS

The sets for analysis will include the Efficacy set, the Modified Efficacy set, the Per Protocol Set, and the Safety set.

4.1. EFFICACY SET

The Efficacy set will consist of all randomized subjects who complete at least 80 days of study drug dosing and have a Week 18 (or ET) visit MRI-PDFF assessment. Subjects in this set will be analyzed according to the treatment to which they were randomized. This set will be utilized for analyses of efficacy.

4.2. MODIFIED EFFICACY SET

The Modified Efficacy set will consist of all randomized subjects who receive at least one dose of study drug and who have at least one post-dose MRI-PDFF assessment. Subjects in this set will be analyzed according to the treatment to which they were randomized. This set will be utilized for analyses of secondary endpoints, as well as select analyses of the primary endpoint.

4.3. PER PROTOCOL SET

The Per Protocol Set consists of all subjects who receive at least one dose of study treatment and who have a Week 18 (or ET) visit MRI-PDFF assessment completed no more than 14 days following the final dose of study treatment.Subjects who have less than 80% treatment compliance will also be excluded from the Per Protocol Set.

4.4. SAFETY SET

The Safety set consists of all subjects who receive at least one dose of study treatment, regardless of whether or not they undergo any study assessments. Subjects in this set will be analyzed according to the treatment which they receive. All safety analyses will be based on this set.

5. STATISTICAL METHODS

Descriptive statistical methods will be used to summarize the data from this study, with hypothesis testing performed for the primary and other selected efficacy endpoints. Unless stated otherwise, the term "descriptive statistics" refers to number of subjects (n), mean, median, standard deviation (SD), minimum and maximum for continuous data and frequencies and percentages for categorical data. Additional statistical methods include analysis of covariance (ANCOVA) and logistic regression. All data collected during the study will be included in data listings. Unless otherwise noted, the data will be sorted first by treatment group, subject number, and then by date within each subject number.

The term 'treatment group' refers to all subjects on the same dosing regimen. There will be three treatment groups in this study:

HTD1801 500 mg BID

HTD1801 1000 mg BID

Placebo BID

All statistical tests will be two-sided with an alpha level of 0.05.

The statistical analyses will be conducted with the SAS® System version 9.4 or higher. All analyses will be subject to formal verification procedures. Specifically, results will be verified utilizing independent programming prior to issuance of the draft statistical report. All documents will be verified by the lead statistician to ensure accuracy and consistency of analyses.

5.1. SUBJECT DISPOSITION, DEMOGRAPHIC AND BASELINE CHARACTERISTICS

Subject disposition will be presented for all randomized subjects. For each treatment group the following will be presented: the number of subjects who meet all eligibility criteria, the number of

subjects included in the Safety, Efficacy and Modified Efficacy sets, the number of subjects who completed the study and discontinued from the study, and the reasons for early discontinuation at any point. Additionally, the number of subjects dosed will be presented and the number of days on study treatment will be summarized for all treated subjects.

Demographic data and baseline characteristics including age, gender, race, and ethnicity, weight at Baseline, height at Baseline, body mass index (BMI), Baseline LFC, Baseline cT1, Baseline alanine aminotransferase (ALT), Baseline aspartate aminotransferase (AST), Baseline ELF, and Baseline HbA1c will be summarized using descriptive statistics for the Safety population, and will be presented by treatment group. This information will be reviewed for baseline differences, but no statistical testing will be performed.

Data on drug accountability and total amount of study drug received will be summarized by treatment group. Total amount of study drug received will be calculated as the sum over the duration of the study of study drug dispensed – study drug returned. Study drug compliance will then be calculated as this amount divided by the expected amount of drug received based on the number of on-treatment days the subject was on the study x 100.

5.2. EFFICACY ANALYSIS

5.2.1. Primary Efficacy Endpoint

The primary endpoint for this study is the absolute change in LFC as measured by MRI-PDFF from Baseline to Week 18.

5.2.2. Primary Efficacy Analysis

LFC values and change from Baseline will be presented by treatment group and study visit. LS means, standard error, and corresponding 95% confidence intervals for absolute change from Baseline in LFC will be presented by treatment group, and LS mean differences between treatment groups will also be presented with their respective standard errors and confidence intervals. Differences in absolute change from Baseline in LFC will be assessed by analysis of covariance that includes treatment group as a fixed effect, and Baseline LFC and Baseline ALT as covariates. Comparisons of each active treatment group to placebo will be tested at the 5% level of significance without adjustment for multiple tests.

5.2.3. Additional Analyses of the Primary Endpoint

As a secondary analysis, the ANCOVA model from the primary analysis, with the addition of Baseline HbA1c group (< 7% vs. ≥ 7%) as a fixed effect, will be used to obtain a comparison of the two active treatment groups, as well as a comparison of the combined active treatment groups vs. placebo, both tested at the 5% alpha level.

5.2.4. Secondary Efficacy Endpoints

- Change from Baseline and percent change from Baseline in fasting glucose from Baseline to Week 18
- Change from Baseline and percent change from Baseline in HbA1c from Baseline to Week 18.
- Proportion of subjects who achieve ≥ 30% relative reduction in LFC as measured by MRI-PDFF from Baseline to Week 18.
- Percent change from Baseline in LFC as measured by MRI-PDFF from Baseline to Week 18.
- Proportion of subjects who normalize LFC to < 5% as measured by MRI-PDFF at Week 18.
- Proportion of subjects who achieve ≥ 5% absolute reduction in LFC as measured by MRI-PDFF from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in homeostasis model assessment-estimated insulin resistance (HOMA-IR) from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in low-density lipoprotein cholesterol (LDL-c) from Baseline to Week 18.

- Change from Baseline and percent change from Baseline in serum triglycerides from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in high-density lipoprotein cholesterol (HDL-c) from Baseline to Week 18.
- Changes from Baseline and percent change from Baseline in AST from Baseline to Week 18.
- Changes from Baseline and percent change from Baseline in ALT from Baseline to Week 18.
- Changes from Baseline and percent change from Baseline in GGT from Baseline to Week 18.
- Proportion of subjects with elevated ALT at Baseline who normalize ALT at Week 18.
- Proportion of subjects with elevated AST at Baseline who normalize AST at Week 18.
- Change from Baseline and percent change from Baseline in Pro-C3 from Baseline to Week 18 for subjects with elevated Pro-C3 at Baseline.
- Change from Baseline and percent change from Baseline in the enhanced liver fibrosis (ELF) score and each component of ELF (tissue inhibitor of metalloproteinases 1 [TIMP-1], N-terminal pro-peptide of type III collagen [PIIINP], and hyaluronic acid [HA]) from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in total bile acid from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in total primary bile acids (and metabolites) from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in total secondary bile acids (and metabolites) from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in total ursodeoxycholic acid (and metabolites) from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in 7α-hydroxy-4-cholesten-3-one (C4) from Baseline to Week 18.
- Change from Baseline and percent change from Baseline in fibroblast growth factor 19 (FGF19) from Baseline to Week 18.

5.2.5. Secondary Efficacy Analysis

Change and percent change from Baseline in LFC will be summarized by treatment group and study visit. Differences in percent change from Baseline in LFC will be assessed by analysis of covariance that includes the effects of treatment group, Baseline LFC, and Baseline ALT, comparing each active treatment group to placebo, as well as a comparison of the combined active treatment groups vs. placebo. A figure summarizing LS means of percent change from Baseline at Week 18 by treatment group will also be generated.

Changes from Baseline to Week 18 in the laboratory assessments listed as secondary efficacy endpoints (fasting glucose, HbA1c, HOMA-IR, LDL-c, serum triglycerides, HDL-c, AST, ALT,

Pro-C3, ELF, total bile acids, C4, FGF19) will be assessed by analysis of covariance that includes the effects of treatment group and Baseline value, comparing each active treatment group to placebo. For these analyses, missing Week 18 values will be imputed using the methods described in section 3.3.4. For HbA1c, this model will also be run without imputed values. For all bile acid analyses other than total bile acid, all analyses will be performed without imputation. Figures summarizing change from Baseline over time by treatment group will also be generated for these laboratory assessments.The number and percentage of subjects who achieve ≥ 30% relative reduction in LFC as measured by MRI-PDFF from Baseline to Week 18 will be summarized by treatment group and study visit. The proportion of subjects who achieve ≥ 30% relative reduction in LFC as measured by MRI-PDFF from Baseline to Week 18 will be analyzed by logistic regression that includes the effects of treatment group, Baseline LFC, and Baseline ALT. A figure summarizing the proportion of subjects achieving ≥ 30% relative reduction in LFC as measured by MRI-PDFF from Baseline to Week 18 by treatment group will also be generated.

The number and percentage of subjects who normalize LFC to < 5% as measured by MRI-PDFF at Week 18 will be summarized by treatment group and study visit. The proportion of subjects who normalize LFC to < 5% as measured by MRI-PDFF at Week 18 will be analyzed by logistic regression that includes the effects of treatment group, Baseline LFC, and Baseline ALT. A figure summarizing the proportion of subjects who normalize LFC to < 5% at Week 18 by treatment group will also be generated.

The number and percentage of subjects who achieve ≥ 5% absolute reduction in LFC as measured by MRI-PDFF from Baseline to Week 18 will be summarized by treatment group and study visit. The proportion of subjects who achieve ≥ 5% absolute reduction in LFC as measured by MRI-PDFF from Baseline to Week 18 will be analyzed by logistic regression that includes the effects of treatment group, Baseline LFC, and Baseline ALT. A figure summarizing the proportion of subjects who achieve ≥ 5% absolute reduction in LFC from Baseline to Week 18 by treatment group will also be generated.

The number and percentage of subjects with elevated ALT at Baseline who normalize ALT at Week 18 will be summarized by treatment group and study visit. The proportion subjects of with elevated ALT at Baseline who normalize ALT at Week 18 will be analyzed by logistic regression that includes the effects of treatment group, Baseline LFC, and Baseline ALT. A figure summarizing the proportion of subjects who normalize ALT at Week 18 by treatment group will also be generated. For this analysis, missing Week 18 values will be imputed using the multiple imputation method as described in section 3.3.4.

If the data is not amenable to analysis by logistic regression for any of the responder endpoints included in this section, then comparisons between treatment groups may be performed using Cochran-Mantel Haenszel (CMH) testing instead.

5.3. SAFETY

5.3.1. Adverse Events

AEs will be mapped to a Medical Dictionary for Regulatory Activities (MedDRA) version 20.1 preferred term (PT) and system organ classification (SOC). If a subject experiences multiple events that map to a single preferred term, the greatest severity grade according to the Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0, and strongest investigator assessment of relation to study medication will be assigned to the preferred term for the appropriate summaries. Should an event have a missing severity or relationship, it will be classified as having the highest severity and/or strongest relationship to study medication. The occurrence of TEAEs will be summarized by treatment group by SOC, PT, and severity. Separate summaries of treatment-emergent serious adverse events (SAEs), TEAEs related to study drug, severe or life-threatening TEAEs, and TEAEs leading to the discontinuation of study treatment will be generated. Additionally, the occurrence of liver-specific AEs will be summarized by treatment group. All adverse events reported will be listed for individual subjects showing verbatim term, PT and SOC. All AEs that occurred prior to the initiation of study treatment will be excluded from the tables but will be included in the listings.

Missing onset dates will be imputed as previously outlined in Section 3.3.5 as required to determine treatment-emergent events.

5.3.2. Clinical Laboratory Assessments, Vital Signs, Height, and Weight

Descriptive summaries of selected (quantitative) clinical laboratory results, and change from Baseline will be presented by time point and treatment group. Laboratory abnormalities will also be summarized by treatment group. Shift tables (low, normal, high) for change from Baseline to Week 18, or last available study visit, will be presented for all appropriate clinical laboratory parameters. Additionally, shift tables (normal, abnormal) for change from Baseline to Week 18, or last available study visit, will be presented for the clinical laboratory parameters listed below. Laboratory abnormalities will be determined using the following criteria:

Laboratory Parameter	**Clinically Significant Abnormal Criteria**
ALP	> 348 U/L
ALT	> 123 U/L
AST	> 102 U/L
Direct Bilirubin	> 0.20 mg/dL
Total Bilirubin	> 2.00 mg/dL

GGT	> 156 U/L in Males > 114 U/L in Females
Glucose	< 60 mg/dL
INR	≥ 1.3
HbA1c	≥ 9.5%

Additionally, descriptive summaries and change from Baseline in weight will be presented by timepoint and treatment group.

Values for all safety variables will be listed by subject and visit (as applicable).

5.3.3. Concomitant Medications

Previous and concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary Enhanced September 2017 version. Previous and concomitant medications will be presented in a data listing. Concomitant medications will additionally be summarized by treatment group.

6. PROTOCOL DEVIATIONS

Possible protocol deviations will be identified and displayed in a data listing and sorted by treatment group, subject and study day (where applicable). Additionally, major protocol deviations will be summarized in a table by type of deviation and treatment group.

7. CHANGES IN THE PLANNED ANALYSES

No deviations in the conduct of the study or the planned analysis are anticipated. Should any deviations from the analyses specified in the authorized statistical analysis plan arise, such deviations will be documented in the final clinical study report.

8. REFERENCES

1. Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics. 1999;55(3):853-857
2. Harrison SA, Rinella ME, Abdelmalek MF, et al. NGM282 for treatment of non-alcoholic steatohepatitis: a multicenter, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet. 2018;391(10126):1174-1185.
3. Loomba R, Sirlin CB, Ang B, et al. Ezetimibe for the treatment of nonalcoholic steatohepatitis: assessment by novel magnetic resonance imaging and magnetic resonance elastography in a randomized trial (MOZART trial). Hepatology. 2015;61(4):1239-1250.
4. Patel NS, Doycheva I, Peterson MR, et al. Effect of weight loss on magnetic resonance imaging estimation of liver fat and volume in patients with nonalcoholic steatohepatitis. Clin Gestroenterol Hepatol. 2015;13(3):561-568.e561.

9. PROGRAMMING CONVENTIONS

- Page orientation, margins, and fonts: Summary tables, listings, and figures will appear in landscape orientation. There should be a minimum of a 1.25" boundary on the upper (bound) edge, and a minimum of a 1.0" boundary on the remaining three edges. Output should be printed in Courier New with a point size of 8. Titles may be printed using a larger font (e.g., Arial point size 10).
- Identification of analysis population: Every summary table and figure should clearly specify the analysis population being summarized. Listings will be prepared for all subjects.
- Group headers: In the summary tables, the group headers will identify the summary group and the sample size for the indicated analysis population. Of note, the header's sample size does not necessarily equal the number of subjects actually summarized within any given summary module; some subjects in the analysis population may have missing values and thus may not be summarized.
- Suppression of percentages corresponding to null categories: When count data are presented as category frequencies and corresponding percentages, the percent should be suppressed when the count is zero in order to draw attention to the non-zero counts.
- Presentation of sample sizes: Summary modules should indicate, in one way or another, the number of subjects actually contributing to the summary statistics presented in any given summary module. As mentioned above, this may be less than the number of subjects in the analysis population due to missing data.
    - In the quantitative modules describing continuous variables (and thus presenting sample size, means, and standard deviations), the sample size should be the number of non-missing observations. The number of missing observations, if any, will be noted.
    - For categorical variables that are presented in frequency tables, the module should present the total count in addition to the count in each category. Percentages should be calculated using this total as the denominator, and the percentage corresponding to the sum itself (that is, 100%) should be presented so as to indicate clearly to a reviewer the method of calculation. The number of missing observations, if any, will be noted.
- Sorting: Listings will be sorted by treatment group, subject number and date, if applicable. If a listing is sorted in a different manner, a footnote will indicate as such.
- General formatting rules: Rounding for all variables will occur only as the last step, immediately prior to presentation in listings, tables, and figures. No intermediate rounding will be performed on derived variables. The standard rounding practice of rounding numbers ending in 0-4 down and numbers ending in 5-9 up will be employed.
- Numerical Values: The presentation of numerical values will adhere to the following guidelines:
    - Raw measurements will be reported to the number of significant digits as captured electronically or on the CRFs.

  - Standard deviations will be reported to one decimal place beyond the number of decimal places the original parameter is presented.
  - Means will be reported to the same number of significant digits as the parameter.
  - Calculated percentages will be reported with no decimals.
- Dates will be formatted as DDMMMYYYY. Partial dates will be presented on data listings as recorded on CRFs.
  - Time will be presented according to the 24-hour clock (HH:MM).

10. PROPOSED TABLES, LISTINGS, AND FIGURES

Summary Tables

Accountability and Baseline Characteristics

14.1.1.1 Subject Disposition – All Randomized Subjects
14.1.1.2 Protocol Deviations – All Randomized Subjects
14.1.2 Demographics and Baseline Characteristics – Safety Set
14.1.3 Extent of Exposure and Treatment Compliace – Safety Set
14.1.4 Concomitant Medications by ATC Classification and Preferred Term- Safety Set

Efficacy

14.2.1.1 Liver Fat Content by Time Point – Efficacy Set
14.2.1.2 Liver Fat Content by Time Point – Modified Efficacy Set
14.2.1.3 Liver Fat Content by Time Point – Per Protocol Set
14.2.2.1 Changes in Fasting Glucose from Baseline to Week 18 – Modified Efficacy Set
14.2.2.2 Changes in HbA1c from Baseline to Week 18 – Modified Efficacy Set
14.2.2.3 Changes in HOMA-IR from Baseline to Week 18 – Modified Efficacy Set
14.2.2.4 Changes in LDL-c from Baseline to Week 18 – Modified Efficacy Set
14.2.2.5 Changes in Serum Triglycerides from Baseline to Week 18 – Modified Efficacy Set
14.2.2.6 Changes in HDL-c from Baseline to Week 18 – Modified Efficacy Set
14.2.2.7 Changes in AST from Baseline to Week 18 – Modified Efficacy Set
14.2.2.8 Changes in ALT from Baseline to Week 18 – Modified Efficacy Set
14.2.2.9 Changes in Pro-C3 from Baseline to Week 18 – Modified Efficacy Set
14.2.2.10 Changes in ELF Score and Components from Baseline to Week 18 – Modified Efficacy Set
14.2.2.11.1 Changes in Total Bile Acids from Baseline to Week 18 – Modified Efficacy Set
14.2.2.11.2 Changes in Total Primary Bile Acids from Baseline to Week 18 – Modified Efficacy Set
14.2.2.11.3 Changes in Total Secondary Bile Acids from Baseline to Week 18 - Modified Efficacy Set
14.2.2.11.4 Changes in Total Ursodeoxycholic Acid from Baseline to Week 18 – Modified Efficacy Set
14.2.2.12 Changes in C4 from Baseline to Week 18 – Modified Efficacy Set
14.2.2.13 Changes in FGF19 from Baseline to Week 18 – Modified Efficacy Set
14.2.2.14 Changes in GGT from Baseline to Week 18 – Modified Efficacy Set

Safety

14.3.1.1 Overall Summary of Treatment-Emergent Adverse Events– Safety Set
14.3.1.2 Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Greatest Severity – Safety Set
14.3.1.3 Treatment-Emergent Adverse Events Related to Study Drug by System Organ

Class and Preferred Term – Safety Set
14.3.1.4 Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term – Safety Set
14.3.1.5 Treatment-Emergent Adverse Events Leading to the Discontinuation of Study by System Organ Class and Preferred Term – Safety Set
14.3.4.1 Clinical Laboratory Chemistry Parameters by Time Point – Safety Set
14.3.4.2 Clinical Laboratory Hematology Parameters by Time Point – Safety Set
14.3.4.3 Clinical Laboratory Urinalysis Parameters by Time Point – Safety Set
14.3.4.4 Clinical Laboratory Abnormalities – Safety Set
14.3.4.5 Clinical Laboratory Parameter Shifts from Baseline to Week 18 – Safety Set
14.3.4.6 Clinical Laboratory Parameter Shifts to Abnormal from Baseline to Week 18 – Safety Set
14.3.5.1 Vital Signs by Study Visit – Safety Set
14.3.5.2 12-Lead ECG Assessments by Study Visit – Safety Set

Summary Figures

Efficacy
14.2.1.3 Bar Plot of Mean Absolute Change from Baseline at Week 18 in LFC
14.2.1.4 Bar Plot of Mean Percent Change from Baseline at Week 18 in LFC
14.2.1.5 Bar Plot of Proportion of Subjects with ≥ 30% Reduction in LFC from Baseline to Week 18
14.2.1.6 Bar Plot of Proportion of Subjects who normalize LFC to < 5% at Week 18
14.2.1.7 Bar Plot of Proportion of Subjects who achieve ≥ 5% absolute reduction in LFC from Baseline to Week 18
14.2.3.1 Bar Plot of Mean Change from Baseline by Time Point in Fasting Glucose
14.2.3.2 Bar Plot of Mean Change from Baseline by Time Point in HbA1c
14.2.3.3 Bar Plot of Mean Change from Baseline by Time Point in HOMA-IR
14.2.3.4 Bar Plot of Mean Change from Baseline by Time Point in AST
14.2.3.5 Bar Plot of Mean Change from Baseline by Time Point in ALT
14.2.3.6 Bar Plot of Proportion of Subjects who Normalize ALT at Week 18
14.2.3.7 Bar Plot of Mean Change from Baseline by Time Point in Lipids
14.2.3.8 Bar Plot of Mean Change from Baseline by Time Point in Pro-C3
14.2.3.9 Bar Plot of Mean Change from Baseline by Time Point in ELF
14.2.3.10 Bar Plot of Mean Change from Baseline by Time Point in Total Bile Acids
14.2.3.11 Bar Plot of Mean Change from Baseline by Time Point in C4
14.2.3.12 Bar Plot of Mean Change from Baseline by Time Point in FGF19

Data Listings

16.2.1.1 Subject Disposition
16.2.1.2 Informed Consent

16.2.2.1 Protocol Deviations/Violations
16.2.2.2 Inclusion/Exclusion Criteria
16.2.3 Reasons for Exclusion from Analysis Sets
16.2.4.1 Demographics
16.2.4.2 Medical History
16.2.4.3 Prior and Concomitant Medications
16.2.5 Study Drug Compliance
16.2.6 Liver Fat
16.2.7.1 Adverse Events
16.2.7.2 Serious Adverse Events
16.2.7.3 Adverse Events Related to Treatment
16.2.7.4 Adverse Events Leading to Discontinuation
16.2.8.1.X Laboratory Tests - Panel

- Hematology and Coagulation
- Chemistry
- Urinalysis
- Bile Acid Panel
- Other

16.2.8.1 Vital Signs, Height, and Weight
16.2.8.2 Pregnancy Test
16.2.9 ECG