Statistical Analysis Plan

A Double-Blind, Randomized, Placebo-Controlled, Phase 2 Study to Explore the Efficacy and Safety of Elobixibat in Adults with Nonalcoholic Fatty Liver Disease (NAFLD) or Nonalcoholic Steatohepatitis (NASH)

PROTOCOL NO.	A3309-012
Sponsor	Elobix AB
CRO	[redacted]
Version	1.0
Date	06-July-2020

SIGNATURE PAGE

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Signature

[redacted]

Date

[redacted]

[redacted]

Signature

[redacted]

Date

DOCUMENT REVISON HISTORY

VERSION	DATE	AUTHOR/ UPDATED BY	COMMENTS
1.0	07/06/2020	NA	NA

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviation	Term
A3309	Elobixibat
AE(s)	adverse event(s)
AESI(s)	adverse event(s) of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
Anti-HCV	hepatitis C antibody
APRI	AST to platelet ratio index
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Class
BMI	body mass index
C4	7α-hydroxy-4-cholesten-3-one
CI	confidence interval
CM	concomitant medication
CRF	case report form
CRP	C-reactive protein
ECG(s)	electrocardiogram(s)
eGFR	estimated glomerular filtration rate
EOT	End of Treatment
EOS	End of Study
FGF-19	fibroblast growth factor-19
FIB-4	fibrosis-4
GCP	Good Clinical Practice
GLP-1	glucagon-like peptide-1
HbA1c	hemoglobin A1c
HDL-C	high-density lipoprotein-cholesterol
ICF(s)	Informed Consent Form(s)
ICH	International Council on Harmonisation
INR	international normalized ratio
IPD	important protocol deviation(s)
IRB	institutional review board

Abbreviation	Term
ITT	intent to treat
IWRS	Interactive Web Response System
LDL	low-density lipoprotein
LDL-C	low density lipoprotein-cholesterol
LS	least squares
lsmean	least squares mean
MCMC	Markov chain Monte-Carlo
MDRD	modification of diet in renal disease
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed models repeated measures
MRI	magnetic resonance imaging
MRI-PDFF	magnetic resonance imaging-proton density fat fraction
NAFLD	nonalcoholic fatty liver disease
NAS	NALFD activity score
NASH	nonalcoholic Steatohepatitis
NFS	NALFD fibrosis score
PCS	potentially clinically significant
PDFF	proton density fat fraction
PP	per protocol
PT	preferred term
QQ Plot	quantile-quantile plot
EM	expectation – maximization
SAE(s)	serious adverse event(s)
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class
TEAE(s)	treatment-emergent adverse event(s)
VLDL	very low-density lipoprotein
WHO	World Health Organization

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report efficacy and safety data for study A3309-012 based on protocol dated on 10 September 2019 (version 3.0).

1.1 STUDY OBJECTIVES

Primary Objective

- To evaluate the efficacy of repeated once-daily doses of 5 mg elobixibat on serum levels of low-density lipoprotein cholesterol (LDL-C) in patients with NAFLD or NASH.

Secondary Objectives

- To assess the safety and tolerability of elobixibat in patients with NAFLD or NASH
- To evaluate the effect of elobixibat on liver steatosis as measured by magnetic resonance imaging (MRI) for (i) liver fat fraction (using proton density fat fraction [PDFF]) and (ii) total liver fat (using whole liver fat volume)
- To evaluate the effect of elobixibat on liver function using alanine aminotransferase (ALT)
- To evaluate the effect of elobixibat on lipids and total bile acids in patients with NAFLD or NASH

1.2 STUDY ENDPOINTS

Primary Efficacy Endpoint

- Change from baseline in serum LDL-C at week 16.

Secondary Efficacy Endpoints

- Absolute change from baseline to week 16 in liver fat fraction (%) as measured by MRI-PDFF
- Absolute change from baseline to week 16 in total liver fat (mL) as measured by MRI
- Change from baseline to week 16, and follow-up in the following:
    - Serum ALT, AST, and gamma-glutamyl transferase
    - HDL-C, non-HDL-C, LDL-C/HDL-C ratio, and triglycerides
    - Total bile acids

Exploratory Efficacy Endpoints

- Change from baseline (unless otherwise specified) to week 4, week 8, week 12, week 16, and follow-up in the following:

- Fasting glucose, insulin, homeostatic model assessment-insulin resistance, HbA1c, free fatty acids
- GLP-1, FGF-19, and C4 levels
- High-sensitivity C-reactive protein (CRP)
- Lanosterol and beta-sitosterol
- Body weight, BMI, waist circumference, and waist-to-hip ratio
- Apolipoprotein A1 and Apolipoprotein B
- Fibrosis and inflammation markers: FIB-4, APRI, and NFS

Safety Endpoints

The safety endpoints for this study include clinical laboratory assessments, vital signs, AEs, ECGs, and concomitant medications.

1.3 STUDY DESIGN

General Study Design and Plan

This is a double-blind, randomized, placebo-controlled, Phase 2 study to explore the efficacy and safety of elobixibat (oral dose of 5 mg) once daily for 16 weeks in patients with NAFLD or NASH. The study includes a 6-week Screening Period, followed by a 16-week Treatment Period, and a follow-up visit 2 weeks after the last dose of study drug.

After completing the Screening period, eligible patients will be randomized on Day 1 (Visit 2) in a 1:1 fashion to either 5 mg/day elobixibat or matching placebo. There will be a total of 7 scheduled visits during the study, as follows:

- Visit 1: Screening period (week -6 through Day -1)
- Visit 2: Day 1/Randomization visit
- Visit 3: Week 4
- Visit 4: Week 8
- Visit 5: Week 12
- Visit 6: Week 16/End of Treatment (EOT). Any patient that discontinues treatment prematurely should complete this visit at the time of discontinuation
- Visit 7: Follow-Up period. The follow-up visit will occur 2 weeks after the last dose of study drug whether the patient completes the study or discontinues prematurely

Additional unscheduled visits may be required for patients who need direct site assistance (e.g., due to AE monitoring, in order to fulfill screening requirements, and/or for safety maintenance).

Randomization and Blinding

A double-blind design is employed so that both the investigators and the patients will be unaware of the treatment assignment during the study. Eligible patients will be randomly assigned in a 1:1 ratio to the study drug or the matching placebo by the Interactive Web Response System (IWRS). Patients who withdraw from the study after the randomization visit are not to be replaced and their randomization number will not be reused.

To ensure blinding, the study drug and the matching placebo have the same shape and size. Labels on the treatment bottles will not identify the treatment to which a patient has been randomized. Traceability of the treatment is ensured by the bottle number. Other than designated randomization personnel, all study conduct persons, investigators, and patients are blinded to study drug assignment.

Per the protocol, total bile acids results are blinded. Operationally, blinding was implemented by not having total bile acids, FGF-19, C4, and lanosterol and beta-sitosterol results available until after database lock. The same blinding approach was applied to the imaging results for the post-baseline imaging results.

Sample Size

Due to the exploratory nature of this study, no formal power calculations were used to determine the sample size. The number of subjects (46 patients in total, 23 patients per treatment group) was chosen based on clinical experience with other similar proof-of- concept studies.

2 STATISTICAL CONSIDERATIONS

2.1 GENERAL CONSIDERATIONS

General principles will be followed:

- Population summaries (e.g., disposition, demographics and baseline characteristics, medical/surgical history, prior and concomitant medications) and efficacy tables will summarize data by the randomized treatment group. Safety tables will be summarized by the actual treatment subject received.
- All continuous variables will be summarized using descriptive statistics including the number of patients with data to be summarized (n), mean, standard deviation (SD), median, minimum (min) and maximum (max). Minimum and maximum will be displayed with a maximum of 2 decimal places. Mean and median will be rounded to 1 additional decimal place, SD will be rounded to 2 additional decimal places, compared to the displayed minimum and maximum.
- Categorical variables will be summarized with count and percent; the denominator will be based on the number of subjects in the treatment group, unless otherwise stated.
- In general, efficacy analysis will be performed using the intent-to-treat (ITT) population and safety analyses will be performed using the safety population. Per-protocol population will be used for the primary endpoint sensitivity analysis.

- The default significance level will be (5%); confidence intervals (CIs) will be 95% and all tests will be two-sided, unless otherwise specified in the description of the analyses.
- P-values will be rounded to three decimal places. If a p-value is less than 0.001, it will be reported as "<0.001." If a p-value is greater than 0.999 it will be reported as ">0.999".

All analyses will be performed using SAS®, Version 9.4 or higher.

2.2 DEFINITIONS OF ANALYSIS SETS

ITT Population

The intent-to-treat (ITT) population will include all patients who are randomized in the study and receive at least 1 dose of study drug. The ITT population will be used for all efficacy analyses based on the treatment randomized unless otherwise specified.

Per Protocol Population

The per-protocol (PP) population will include all ITT patients who finish Visit 6/EOT with valid LDL-C measurements and do not have any important protocol deviations that may impact the analysis of efficacy. The important protocol deviations are:

- Failure to meet the Inclusion criteria
- Failure to meet the Exclusion criteria
- Compliance rate of <80% or >120% with study drug administration
- Took disallowed meds over 50% time of treatment period
- Others (mis-randomized, took the wrong drug, etc.)

Safety Population

The safety population will include all patients who are randomized and receive at least 1 dose of study drug and will be used for all safety analyses based on the treatment received. The Safety and ITT Populations are the same, as there were no subjects that did not receive the treatment to which they were randomized.

2.3 CONVENTIONS FOR DATA HANDLING

Baseline

Baseline for a particular variable will be defined as the last available assessment for that variable prior to the initiation of first dose of study drug.

Study Day

Study Day will be calculated from the reference start date and will be used to show start/stop day of assessments and events.

Reference start date (Study Day 1) is defined as the day of first dose for efficacy and safety analysis. Note: A few patients were randomized but not treated until the next day.

For assessments or events after the reference start date

Study day = date of assessment or event – reference start date + 1

For assessments or events prior to the reference start date

Study day = date of assessment or event – reference start date

Analysis Visit

Visit-based data summaries will be presented by analysis visits. A window assessment will be performed according to the below table for any subject that missed a planned visit (on both efficacy and safety endpoints). The purpose of using this analysis window is to include potential unscheduled visits when a subject missed the planned visit. If a subject's planned visit assessment is out of the defined visit window and there is no unscheduled assessment within the same window, then this CRF collected planned visit assessment will be used for the analysis.

Period	Analysis Visit	Protocol Defined Window	Scheduled Study Day	Windows
Screening	Visit 1 (Screening)	-6 Weeks to Day -1	-41 to -1	<= -1
Treatment	Visit 2/ Randomization (Day 1)	Day 1	1	1
	Visit 3 (Week 4)	Week 4 (±3 days)	28	2, 42
	Visit 4 (Week 8)	Week 8 (±3 days)	56	43, 70
	Visit 5 (Week 12)	Week 12 (±3 days)	84	71, 98
	Visit 6/EOT (Week 16)	Week 16 (±3 days)	112	99, 119
Follow-up	Visit 7 (Follow-up)	Week 18 (±3 days)	126	>=120

If multiple records in the same analysis visit, the scheduled visit is preferred to the unscheduled visits, otherwise the one closest to the scheduled visit study day will be selected for the analysis and summaries. The later one is preferred if there are more than one record have the same distance to the scheduled study day.

In the derivation of post baseline extreme values (e.g., minimum, maximum, worst, last), all post baseline assessments will be included.

All data including additional unscheduled visits required for patients who need direct site assistance (due to AE monitoring, in order to fulfill screening requirements, and/or for safety maintenance) will be presented in subject listings unless otherwise specified.

Missing and Partial Date

No imputations will be done for study visit dates, birth dates, informed consent date, and assessment dates.

Partial concomitant medication, and AE start dates will be imputed as following:

- If missing day, use the later of (01 of the month, the first dose date) if the first dose is at the same year and month of AE/CM, otherwise use the 01 of the month.
- If missing day and month, use the later of (01JAN of the year, first dose date) if the first dose year is at the same year of AE/CM, otherwise use the 01JAN of the year.
- If completely missing, no imputation will be applied

Partial concomitant medication and AE end dates will be imputed as the following:

- If missing day, use the last day of the month
- If missing day and month, use 31st December
- If completely missing, no imputation will be applied

Partial initial diagnostic dates will be imputed as: if day and/or month are missing use 01 and/or Jan. If year is missing, then no imputation, treat the date as missing. Any other imputation will be based on evidence from additional data, and will be noted for each case, should it occur.

2.4 CHANGES TO THE STATISTICAL ANALYSIS FROM THE PROTOCOL

Fibrosis and inflammation markers (FIB-4, APRI, and NALFD fibrosis score (NFS)) have been added as exploratory endpoints. Additional analyses may include analyses of subjects stratifying by baseline Fib-4 at 1.3. Pro-C3 results, indicating fibrosis, may also be available for analysis.

Physical examination (including general appearance, eyes, nose, throat, head/neck/thyroid, lymph nodes, cardiovascular, lungs/chest, abdomen, genitourinary, extremities, skin, musculoskeletal, neurologic, and other) as listed in protocol section 10.2.4 are not collected.

2.5 IMPACT OF COVID-19 ON THE CLINICAL TRIAL

During the data review meeting in which major protocol violations are identified, violations of the protocol that were the result of the COVID-19 pandemic will be noted.

3 STATISTICAL ANALYSES

3.1 SUBJECT INFORMATION

Disposition of Subjects

The number and percentage of subjects for the following will be summarized by randomized treatment group and overall where applicable:

- Patients enrolled (who signed the informed consent)
- Patients included in each analysis set

- Patients completed the study
- Patients discontinued treatment early (including EOT reason)
- Patients discontinued study early (including EOS reason)

The disposition listing will include information such as: included in an analysis set, first dose date, end of treatment (EOT) status, end of study (EOS) status, and with EOT/EOS reasons.

Subjects that were screen failures and the reason for screen fail will be provided in a listing.

Per Version 3.0 of the protocol, subjects that are screen failures, may be rescreened if approved by the Medical Monitor. Subjects that have been rescreened will be noted.

Protocol Deviations and Violation of Inclusion/Exclusion Criteria

All personnel involved with the analysis of the study will remain blinded until database lock for the double-blind treatment period and until major protocol violations have been identified.

Subjects who received study drug but violated any inclusion or exclusion criteria will be listed.

Protocol deviations will be captured for each site and reviewed by the medical monitor to identify major protocol deviations from the list. If noted stated elsewhere, a major protocol violation is a deviation that has an impact on patient safety, may substantially alter risks to patients, may have an effect on the integrity of the study data, or may affect the patient's willingness to participate in the study.

All major protocol deviations will be presented in a listing.

Demographics and Baseline Characteristics

Descriptive summaries of demographics and other baseline characteristics will be presented by treatment group and overall using the Safety/ITT population. These summaries will include age, gender, childbearing potential, race, ethnicity, height, weight and body mass index (BMI).

Conversion factors and calculations for height and weight are:

- Height (in cm) = height (in inches) * 2.54
- Weight (in kg) = weight (in lbs) * 0.4536

By subject listings for demography and baseline characteristics will be provided.

Medical and Surgical History

NASH/NAFLD history variables characteristics will be summarized by randomized treatment group and total on ITT.

Medical histories include type of disease under study (NASH or NAFLD), duration of disease under study, NASH confirmed by live biopsy (Yes or No), and whether patients have any other relevant medical or surgical history. The duration of disease under study is calculated as (the date of randomization – date of initial diagnosis + 1)/365.25. Imputed date of initial diagnosis date will be used if the collected date is partial.

If applicable, the other relevant medical and surgical history will be coded by latest MedDRA version. The number (percent) of subjects reporting a history of any other relevant medical or surgical history, will be summarized for each randomized treatment group and overall.

A subject data listing of medical and surgical history will be provided.

Prior and Concomitant Medications

All medications with a stop date prior to and within 3 months (92 days) of the date of first dose are regarded as prior medication. All medications with a start date or stop date on or after the first dose date are regarded as concomitant medication. Medications where start date and end date are both completely missing (both dates cannot be imputed) will be treated as concomitant medications.

The number (percent) of subjects who took prior and concomitant medications will be summarized respectively by Anatomic Therapeutic Chemical (ATC) level and preferred term for each randomized treatment group and overall on ITT. Subjects taking the same medication multiple times will only be counted once for that ATC class and preferred term. A subject level listing will also be presented.

All investigator terms for medications recorded on the CRF will be coded using the World Health Organization (WHO) Drug Dictionary B3, March 2019 version.

3.2 EFFICACY ANALYSES

Primary Endpoint

The primary efficacy endpoint for this study is the change from baseline in serum LDL-C at week 16.

Change from baseline in serum LDL-C will be analyzed by using an analysis of covariance (ANCOVA) after the missing data imputation. The ANCOVA model will include treatment group as a factor, baseline LDL-C and baseline LDL-C by treatment interaction values, and baseline lipid lowering medications (Yes or No) as covariates.

Serum LDL-C results and the change from baseline at each scheduled time point, along with the minimum and maximum post baseline value, will be summarized using descriptive statistics.

Box plots for serum LDL-C results absolute value and change from baseline overtime will be provided.

Multiple Imputation for Primary Endpoint Variable

If a subject terminates the study prematurely, the subject's LDL-C values may not be collected after termination. Since the estimand of this study is treatment difference at week 16 in the ITT population, a multiple imputation procedure (

Angulo, J. M. et al. (2007), "The NAFLD fibrosis score: A noninvasive system that identifies liver fibrosis in patients with NAFLD," Hepatology, 45, 846-854.

utilizing the regression method will be used to impute each missing data point up to week 16 with a set of plausible values that represent the uncertainty about the values to be imputed. The LDL-

C data from completers in the placebo arm will be used to impute the missing data in early terminators. *The imputed values should have an LDL-C value that is greater than minimum of all observed values minus 15% or less than maximum of all observed values plus 15%.* The following options in SAS Proc MI statement will be used to limit the minimum and maximum imputed values: Minimum=x and Maximum=y, where x and y are the minimum and maximum values stated above. With this option, when the imputed value is outside of this limit, the MI procedure redraws another value until the imputed value is within the limits.

The steps below will be followed in the imputation:

- **Step 1:** Non-monotone missing data at intermediate visits will be imputed for *200* times to generate 200 imputed datasets, so the data will have a monotone missing pattern. SAS® procedure *PROC MI* will be used for the multiple imputation. The monotone Markov Chain Monte Carlo (MCMC) method will be used at this step. Non-convergence may occur at this step. If the MCMC imputation does not converge, additional imputations will be repeated. The first 200 convergent MCMC imputations will be used in the analysis.

  In the SAS MI procedure at this step, the following options will be used: nimpute = 200 and seed = 396858. The following options will be used in the MCMC statement: impute = monotone, chain = multiple, initial = EM, prior = Jeffreys. In the EM statement, the option maxiter=200 will be used. The var statement will include treatment arm, baseline LDL-C value, baseline lipid lowering medications, and the change from baseline in LDL-C value at each scheduled post-randomization visit up to week 16. Sample SAS codes for multiple imputations are provided in the Appendix.

- **Step 2:** After the intermediate missing values have been imputed, subjects are categorized to 3 groups: 1) Completers (defined as subjects who completed the study treatment) in the control arm with non-missing value at week 16. 2) Subjects with a missing value at week 16 in both arms. Missing values in these subjects will be imputed using subjects in group 1. 3) Completers in the active arm with non-missing value at week 16 (based on the analysis visit window in section 2.3). These subjects will not be used to impute missing values in the next step.

- **Step 3:** At this point, each subject in the 200 data sets generated at step 1 either has complete data or has monotone missing data. Subjects in group 1 in Step 2 will be used to impute missing values in subjects in group 2. Missing data in group 2 will be imputed as the predicted value from a regression model with treatment arm, LDL-C baseline, baseline lipid lowering medications, and the change from baseline of all previous scheduled visits included in the model. Since the dataset has a monotone missing data pattern, the imputation process will start from the earliest visit with missing data (e.g. week 4) through week 16 in a sequential order. SAS Proc MI with option Monotone Reg, nimpute=1, and seed = 25380 will be applied. The imputation will be done for each one of the *200* data sets generated at step 1.

- **Step 4:** After imputation for monotone missing values at Step 3, subjects in all 3 groups will be pooled and sorted by imputation sequence (variable name _Imputation_ from Step 1).

- **Step 5:** Each imputed complete dataset will be analyzed using the analysis of covariance (ANCOVA) model with the randomized treatment group as a factor and baseline LDL-C values

as a covariate. The estimates of least-square means (lsmeans), least-square mean differences 95% CI, and p-value will be generated for each one of the 200 imputed complete datasets.

- **Step 6:** The results from the 200 complete datasets, including lsmeans and lsmean difference with their corresponding 95% confidence intervals, and the p-value between elobixibat 5 mg and placebo will be combined using SAS® procedure PROC MIANALYZE for statistical inference.

The above primary efficacy analysis will be repeated using the PP population.

Sensitivity Analysis for primary endpoint

Mixed models repeated measures (MMRM) analysis based on both imputed data and observed data will be used as sensitivity analysis. The explanatory variables will include treatment, week, baseline value, baseline lipid lowering medications, treatment-by-week interaction, and treatment-by-baseline interaction, with unstructured covariance model across week. The average treatment effect at Week 16 (lsmeans with 95% CI for each treatment group and lsmean difference with 95% CI) will be provided as both tables and figures. Additional MMRM analyses may include the addition of other model covariates such as age at baseline, sex, and baseline FIB-4 (dichotomized at ≥1.3). An additional series of models may include percent change from baseline as the dependent variable.

A plot of fitted values versus residuals and the residual QQ Plot will be used to check the ANCOVA model assumptions. If there are concerns about the model assumptions, i.e., normality and homogeneity of variances, a non-parametric ANCOVA based on the rank scores will be used as a sensitivity analysis as well.

Secondary Efficacy Endpoints

A summary of descriptive statistics of the absolute change from baseline to week 16 in liver fat fraction (%) as measured by MRI-PDFF and total liver fat (mL) as measured by MRI will be provided.

The test results and changes from baseline at each scheduled time point, including week 16 and Follow-up, will be summarized for the following variables:

- Serum ALT, AST, and gamma-glutamyl transferase
- HDL-C, non-HDL-C, LDL-C/HDL-C ratio, and triglycerides
- Total bile acids

The minimum and maximum post baseline value for each above lab test will be included in the visit summaries.

The change from baseline at week 16 will be analyzed using the analysis of covariance (ANCOVA) with the treatment group as a factor and the corresponding baseline value as a covariate. The model assumptions for each secondary efficacy endpoint will be evaluated before the analysis. If there are concerns about the model assumptions, i.e., normality and homogeneity of variances, non-

parametric ANCOVA based on the rank scores will be used. No missing data imputation for secondary efficacy endpoints.

MMRM figures of change from baseline by study arm and the difference between study arms will be generated for all secondary endpoints.

Exploratory Efficacy Endpoints

The test results and change from baseline (unless otherwise specified) to week 4, week 8, week 12, week 16, and follow-up, will be summarized for the following variables:

- Fasting glucose, insulin, homeostatic model assessment-insulin resistance, HbA1c, free fatty acids
- GLP-1, FGF-19, and C4 levels
- High-sensitivity CRP
- Lanosterol and beta-sitosterol
- Body weight, BMI, waist circumference, and waist-to-hip ratio (Physical examination)
- Apolipoprotein A1 and Apolipoprotein B
- Fibrosis and inflammation markers: FIB-4, APRI and NFS

By patient figures (spaghetti plots) plots for the absolute value and change from baseline over time will be provided for each efficacy endpoint. An MMRM of change from baseline will be provided for all efficacy endpoints. Two plots for each MMRM analysis will be provided, one figure by treatment arm, and a second figure for the contrast between treatment arms. Correlation matrices will be provided for efficacy endpoints overall, by treatment, and by treatment by study visit.

FIB-4 = [Age (years) × AST (U/l)]/[Platelet($\times 10^9$/l) × (ALT(U/l))$^{0.5}$]

APRI = [[AST (IU/l)/AST (ULN)]/ Platelet($\times 10^9$/l)] × 100

NAFLD fibrosis score = −1.675 + 0.037 × age (years) + 0.094 × BMI (kg/m^2) + 1.13 × IFG/diabetes (yes = 1, no = 0) + 0.99 × AST/ALT ratio − 0.013 × platelet ($\times 10^9$/l) − 0.66 × albumin (g/dl)

Additional analyses may be performed to identify differences in treatment effect by baseline characteristics, such as dichotomized subjects by FIB-4 <1.3, and ≥1.3, BMI (<30/≥30), type 2 diabetes (Y/N), cardiovascular co-morbidities (Y/N), lipid lowering concomitant medications (Y/N), and LDL-C (by < /≥ median value).

3.3 STUDY DRUG EXPOSURE AND TREATMENT COMPLIANCE

Exposure will be analyzed by calculating the number of days with exposure to study drug. Descriptive summary statistics will be provided using the safety population.

The treatment compliance will be calculated as (the actual total dose / the expected total dose) ×100, where:

- The actual total dose = (total number tablets of dispensed – total number of tablets returned) × 5 (Note: If tablets were not returned, the data will be considered missing.)
- The expected total dose = (last dose date – first dose date +1) × 5

The number of patients with an overall study treatment compliance <80%, between 80% and 120%, and >120% will be presented based on the safety population.

Treatment compliance for below time periods will be calculated and presented as the same categories as overall compliance:

- Day 1 to Week 4 [study day 1 to study day 28]
- Week 4 to Week 8 [study day 29 to study day 56]
- Week 8 to Week 12 [study day 57 to study day 84]
- Week 12 to Week16 [study day 85 to study day 112]

A subject level listing for study drug exposure (the first dose date, last dose date, duration of treatment) and compliance (including overall compliance and compliance by each time period defined above) will be presented.

Study drug dispensing and accountability will be listed as well.

3.4 SAFETY ANALYSIS

All safety analyses will be performed on the safety population. No inferential statistics will be performed; only summary statistics will be provided unless otherwise noted. Safety data will be presented by actual treatment group.

Safety variables include AEs, vital signs, 12-lead ECG, and clinical laboratory results.

Adverse Events

Adverse events will be coded to primary System Organ Class (SOC) and preferred term (PT) using MedDRA Version (Version 22.0).

TEAEs will be defined as AEs that are new or worsening after the first dose of study drug AEs with missing onset and end date will be considered as a TEAE. . Serious adverse events (SAEs) are captured in the CRF and are determined by the investigator based on protocol section 10.1.1.2. Adverse events of special interest (AESIs) are recorded in the CRF and are defined in protocol section 10.1.1.3.

A general summary or patients with TEAEs, SAEs and AESIs will be tabulated with numbers and percentages of patients per treatment group. The number of AEs leading to withdrawal and SAEs leading to death will also be summarized.

The incidence of TEAEs will be reported as the number (percent) of subjects with TEAEs within SOC and PT. Subjects will be counted only once within a SOC and PT, even if the subject

experienced more than one TEAE within a specific SOC and PT. The number (percent) of subjects with TEAEs will also be summarized by severity (mild, moderate, or severe) and relationship (unrelated, unlikely related, possibly related, probably related, or definitely related) as captured in the CRF. The relationship of each AE to the study drug will be grouped as related (unlikely related, possibly related, probably related, or definitely related, missing) or unrelated.

The following incidence summary tables will be provided:

- All TEAEs
- All TEAEs by severity
- All TEAEs by relationship
- AEs with outcome of death
- Serious TEAEs
- Serious TEAEs related to study drug
- AEs leading to discontinuation of treatment
- All treatment emergent AESI

By subject listings for all AEs (including TEAE and non-TEAE), serious TEAE, treatment emergent AESI, related TEAE, AEs leading to study drug discontinuation, and AEs with an outcome of death will be provided.

Vital Signs

Evaluation of vital signs (systolic and diastolic blood pressure, pulse, respiratory rate and body temperature) will be performed at all visits. Vital signs will be summarized as actual value and change from baseline for each treatment group at each scheduled assessment.

A listing of all vital sign assessments will be presented.

12-Lead ECG

ECG assessments are performed at the Screening visit and the EOT visits. Descriptive statistics for ECG parameters (Heart rate, PR Interval, QRS Interval, QT Interval and QTC Interval) and changes from baseline at each scheduled timepoint will be presented by treatment group.

A subject data listing will present changes from baseline in ECG results captured in the CRF (categorized as normal; abnormal, not clinically significant; and abnormal, clinically significant) to end of treatment.

Laboratory Tests

Descriptive summaries of laboratory data (for clinical chemistry, hematology/coagulation, urinalysis, lipid profile, endocrinology, bile acids, serology, and other tests as listed in the Appendix) will be presented for the safety analysis set along with listings of lab data collected. Laboratory results at each scheduled visit will be summarized by actual treatment group, along

with changes from baseline to each post-baseline visit. Values outside of the reference range will be flagged in the listings.

Shift tables for the shift from baseline with respect to the reference normal range will be provided.

Laboratory abnormalities of special interest will be summarized.

Pregnancy Test

For all females of child-bearing potential, a listing of pregnancy test including test date, test result and reason if test not taken will be presented.

4 MINIMIZING RISK TO TRIAL INTEGRITY DURING THE COVID-19 PANDEMIC

Missed or delayed study visits as a result of the COVID-19 pandemic have occurred. A listing of missed study visits and procedures will be provided. Should the MRI not occur on schedule, that is at the last visit on treatment, the validity of the liver fat quantification (measured by MRI PDFF) results may be compromised, as the potential effects of study drug on liver fat are expected to decrease once study treatment is stopped, although the speed of this loss of potential response is unknown. This loss of documented benefit may also be seen with other efficacy-based laboratory tests. For this reason, these endpoints may be tested in the per protocol population, in addition to the ITT population.

REFERENCES

Angulo, J. M. et al. (2007), "The NAFLD fibrosis score: A noninvasive system that identifies liver fibrosis in patients with NAFLD," Hepatology, 45, 846-854.

Barnard, J. and Rubin, D.B. (1999), "Small-Sample Degrees of Freedom with Multiple Imputation," Biometrika, 86, 948–955.

FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency Guidance for Industry, Investigators, and Institutional Review Boards March 2020 Updated on May 14, 2020.

Lin ZH, Xin YN, Dong QJ, et al. (2011), "Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: an updated meta-analysis. Hepatology," 53, 726-36.

Rubin, D.B. (1987). Multiple Imputation for Nonresponse in Surveys, John Wiley & Sons, Inc.

5 STERLING RK, LISSEN E, CLUMECK N, ET. AL. (2006), "DEVELOPMENT OF A SIMPLE NONINVASIVE INDEX TO PREDICT SIGNIFICANT FIBROSIS PATIENTS WITH HIV/HCV CO-INFECTION," HEPATOLOGY, 43, 1317-1325.APPENDIX

Schedule of Assessments

VISIT	SCREENING PERIOD	TREATMENT PERIOD					FOLLOW-UP PERIOD
	VISIT 1	VISIT 2/ RANDOMIZA-TION	VISIT 3	VISIT 4	VISIT 5	VISIT 6/ END OF TREATMENT	VISIT 7
STUDY WEEK	WEEK -6 TO DAY -1	DAY 1	WEEK 4	WEEK 8	WEEK 12	WEEK 16	WEEK 18
VISIT WINDOW			±3 DAYS	±3 DAYS	±3 DAYS	±3 DAYS	±3 DAYS
Informed consent Medical/surgical history[f] Demographics Serology[a] Review alcohol consumption[j]	X						
Inclusion/exclusion criteria	X	X					
Randomization[g]		X					
12-lead ECG	X					X	
Physical examination[b]	X	X				X	
Vital signs[c] Pregnancy test[d] Clinical safety laboratory tests[h] GLP-1, FGF-19, C4[h] Total bile acids[h] Lipid profile[h] Store blood samples[i]	X	X	X	X	X	X	X
MRI[e]	X					X	
Endocrinology[h]	X	X	X			X	
Fibrosis and inflammation markers[h]		X	X			X	
High-sensitivity CRP	X		X			X	
Dispense study drug		X	X	X	X		

Review concomitant medications and adverse events	X	X	X	X	X	X	X

[a] Includes hepatitis B virus surface antigen and hepatitis C virus antibody and RNA

[b] Includes height (at Screening only), weight, waist circumference, and waist-to-hip ratio. Body mass index will be calculated based on height and weight. A full physical examination will be performed at Screening and at Visits 2 and 6.

[c] Vital signs include blood pressure, heart rate, respiratory rate, and oral body temperature.

[d] For women of childbearing potential, a serum pregnancy test will be performed at Screening, and a urine pregnancy test will be performed at each visit at which study drug is dispensed.

[e] Patients will have MRI to measure liver fat (PDFF), and total liver volume. Patients who discontinue before Visit 6 (Week 16) should have an MRI performed at End of Treatment if they completed at least 4 weeks of treatment.

[f] Liver biopsy results for confirmed NASH within 6 months of Screening to confirm patient eligibility will be collected in the eCRF, if available. [g] Randomization will occur at Visit 2 to assign patients to either 5 mg/day elobixibat or placebo in 1:1 ratio.

[h] Laboratory Parameters. Clinical safety laboratory tests include clinical chemistry, hematology/coagulation, and urinalysis [i]

[j] History of alcohol consumption will be obtained at Screening.

Abbreviations: C4: 7α-hydroxy-4-cholesten-3-one; CRF: case report form; CRP: C-reactive protein; ECG: electrocardiogram; EOT: End of Treatment; FGF-19: fibroblast growth factor-19; GLP-1: glucagon-like peptide-1; INR: international normalized ratio; MRI: magnetic resonance imaging; NASH: nonalcoholic steatohepatitis; PDFF: proton density fat fraction

Laboratory Parameters

CLINICAL CHEMISTRY (FASTING)	HEMATOLOGY/ COAGULATION	URINALYSIS
Albumin	Hematocrit	Bilirubin
ALT [b]	Hemoglobin	Blood
Alkaline phosphatase	Platelet count	Glucose
Amylase	Red blood cell count	Ketones
AST [b]	White blood cell count	Leukocytes esterase
Bicarbonate	Neutrophils	Nitrites
Blood urea nitrogen	Eosinophils	Microscopy including leukocytes (if applicable)
Calcium	Basophils	pH
Chloride	Lymphocytes	Protein
Creatine kinase	monocytes	Specific gravity
eGFR by MDRD (screening only)	International normalized ratio (INR)	Urobilinogen
Gamma-glutamyl transferase [b]	Prothrombin time	
Glucose [c]		
Hemoglobin HbA1c [c]		
Inorganic phosphorus		
Lactate dehydrogenase		
Lipase		
Potassium		
Sodium		
Total bilirubin		
Total protein		
Uric acid		
LIPID PROFILE (FASTING)	**OTHER TESTS**	**ENDOCRINOLOGY**
Apolipoprotein A1 [c]	High-sensitivity C-reactive protein (CRP) [c]	Follicle-stimulating hormone[2]
Apolipoprotein B [c]	C4 [c]	Homeostatic model assessment-insulin resistance [c]
Chylomicron cholesterol	Fibroblast growth factor-19 (FGF-19) [c]	Insulin [c]
Chylomicron triglycerides [b]	FIB-4	Thyroxine[3]
High-density lipoprotein (HDL) cholesterol [b]	APRI	Triiodothyronine[3]
Lipoprotein(a) cholesterol [b]	NAS	Thyroid-stimulating hormone[3]
Low-density lipoprotein (LDL) cholesterol [1][a]	Glucagon-like peptide-1 (GLP-1) [c]	Free fatty acids [c]
LDL triglyceride [b]		Adipose tissue insulin resistance

Non-high-density lipoprotein cholesterol [b]		Adiponectin
Total cholesterol		
Triglycerides [b]		
Very low-density lipoprotein (VLDL) cholesterol		
VLDL triglyceride		
Lanosterol [c]		
Beta-sitosterol [c]		
BILE ACIDS (FASTING)	**PREGNANCY**	**SEROLOGY**
Total bile acids [b]	Serum	Hepatitis B virus surface antigen
	Urine	Hepatitis C virus antibody and ribonucleic acid

1. LDL-cholesterol will be determined by calculation when triglyceride levels are less than or equal to 400 mg/dL, LDL-C will be determined by preparative ultracentrifugation when triglyceride levels are greater than 400 mg/dL.
2. A postmenopausal state is defined as no menses for ≥12 months without an alternative medical cause. A follicle-stimulating hormone level in the postmenopausal range will be used to confirm a postmenopausal state in women ≤55 years of age, performed at Screening visit only.
3. Thyroid function test will be performed during Screening visit only.
[a] Primary efficacy endpoint.
[b] Secondary efficacy endpoints.
[c] Exploratory efficacy endpoints.

SAS sample code for multiple imputation

- For subjects terminated the study prematurely:

/*Step 1 Non-monotone missing data at intermediate visits will be imputed to generate imputed dataset using MCMC method. */

```
PROC MI data=LDL-C out=Mono round=0.1
        seed=396858 nimpute=200;
MCMC IMPUTE=monotone;
VAR  trtp base lipid_low chg1 chg2 chg3 chg_eot;
RUN;
```

Where chg1- chg3 represents the change from baseline at week 4, week 8, and week 12; chg_eot represents the change from baseline at week 16.

/*Step 2: subjects are categorized to the following groups**/

```
data Mono_1 grp3;
        set Mono;
        if trtp='Placebo' and chg_eot^=. then do; group=1; output
Mono_1; end;
        if chg_eot=. and not (trtp='Drug' and completer='Y') then do;
group=2; output Mono_1; end;
        if trtp='Drug' and chg_eot^=. then do; group=3; output grp3; end;
       if trtp='Drug' and completer='Y' and chg_eot=. then do;
group=4; output grp3; end;
run;
```

/*Step3 use group1 and group2 to do imputation*/

```
PROC MI data=Mono_1 out=imputed_1 seed=25380 nimpute=1;
By _imputation_;
CLASS group;
VAR trtp base lipid_low chg1 chg2 chg3 chg_eot;
MONOTONE  reg(chg1=base lipid_low);
MNAR model (chg1/modelobs = (group='1'));
MONOTONE  reg(chg2=base lipid_low chg1);
MNAR model (chg2/modelobs = (group='1'));
MONOTONE reg(chg3=base lipid_low chg1 chg2);
MNAR model (chg3/modelobs = (group='1'));
MONOTONE  reg(chg_eot=base lipid_low chg1 chg2 chg3);
MNAR model (chg_eot/modelobs = (group='1'));
RUN;
```

/*Step 4 After imputation for monotone missing values at Step 3, subjects in all 3 groups will be pooled and sorted */

```
data imputed;
        set imputed_1 grp3;
```

```
proc sort;
   by _Imputation_;
run;
```

/*Step 5: perform ANCOVA for 200 times using imputed datasets*/

```
PROC MIXED data=imputed plots(maxpoints=none);
   by _Imputation_;
        CLASS trtp;
        MODEL chg_eot=trtp base trtp*base;
        LSMEANS trtp/diff cl;
        ODS OUTPUT diffs=diffs lsmeans=lsmeans;
RUN;

 proc sort data=lsmeans;
    by trtp _Imputation_;
 run;
```

/*Step 6: Sample code for combining lsmeans and diffs is as below*/

```
PROC mianalyze data=lsmeans;
BY trtp;
MODELEFFECTS estimate;
STDERR stderr;
RUN;

PROC mianalyze data=diffs;
MODELEFFECTS  estimate;
STDERR stderr;
RUN;
```

- In case there is any missing week-16 serum LDL-C for on-treatment subjects in the active arm, the same imputation method will be performed using the completers' data in the active arm. Note that placebo completers with a missing week-16 LDL-C will be imputed by subjects with non-missing week-16 value in the same arm in steps 2 and 3 above.

/* Step 3.2: Execute the following code there is any missing week-16 serum LDL-C for on-treatment subjects in the active arm. Then, continue to step 4 above. */

```
PROC MI data=grp3 out=grp3 seed=25380 nimpute=1;
By _imputation;
CLASS group;
VAR base lipid_low chg1 chg2 chg3 chg_eot;
MONOTONE reg(chg1= base lipid_low);
MNAR model (chg1/modelobs = (group=' '3'));
```

```
MONOTONE reg(chg2= base lipid_low chg1 );
MNAR model (chg2/modelobs = (group= ' '3'));
MONOTONE reg(chg3= base lipid_low chg1 chg2);
MNAR model (chg3/modelobs = (group= '3'));
MONOTONE  reg(chg_eot= base lipid_low chg1 chg2 chg3);
MNAR model (chg_eot/modelobs = (group= '3'));
RUN;
```

SAS sample code for MMRM

```
proc mixed data=mmrm method=reml;
  class trtp usubjid week lipid_low;
  model chg= trtp base lipid_low trtp*base week trtp*week;
  repeated week/ type = un subject=subnum;
  lsmeans week /pdiff cl;
  ods output lsmeans = lsmeans ;
run;
```

Note: variables included in the model and class statement varies for different endpoints;