MEDPACE

STATISTICAL ANALYSIS PLAN

Protocol Title:	A Phase 2a, Randomized, Double-Blind, Multicenter, Placebo-Controlled Study to Assess the Safety, Tolerability, Pharmacokinetics, and Efficacy of EYP001a in Patients With Nonalcoholic Steatohepatitis
Protocol Number:	EYP001-202
Current Protocol Version/Date:	V5.0 (Amendment 4): 24Mar2020
Investigational Product:	EYP001a
Sponsor:	ENYO Pharma SA Bâtiment B - Bioserra 60 Avenue Rockefeller, 69008 Lyon, France Email: 202@enyopharma.com Telephone: +33 4 37 70 02 19
SAP Version/Date:	V4.0: 02JUL2021

SIGNATURE PAGE

Protocol Title: A Phase 2a, Randomized, Double-Blind, Multicenter, Placebo-Controlled Study to Assess the Safety, Tolerability, Pharmacokinetics, and Efficacy of EYP001a in Patients With Nonalcoholic Steatohepatitis

Protocol Number: EYP001-202

We, the undersigned, have reviewed and approved this Statistical Analysis Plan:

Signature	**Date**
Miriam Anderson Miriam Anderson Senior Statistical Scientist Medpace, Inc.	*Electronically signed by: Miriam Anderson Reason: Approved Date: Jul 2, 2021 14:27 MDT* DD-MMM-YYYY
Ping Xu Ping Xu Director, Biostatistics Medpace, Inc.	*Electronically signed by: Ping Xu Reason: Approved Date: Jul 2, 2021 15:33 CDT* DD-MMM-YYYY
HUGO GIRMA Hugo Girma, MSc Biostatistician, Data Manager ENYO Pharma SA	*Electronically signed by: HUGO GIRMA Reason: Approved Date: Jul 3, 2021 11:05 GMT+2* DD-MMM-YYYY

VERSION HISTORY

Version	Version Date	Description
1.0	29Jan2020	Initial version
2.0	25May2021	Updated for Part B of the study, after Part A is complete.
3.0	01Jul2021	• Clarify: Patients with missing baseline cT1 will not be included in the analysis. • Add summary of derived pruritus TEAEs
4.0	02Jul2021	In Section 3.4.2.5 for Lipoprotein Profiles, add percent change from baseline. Also change ANCOVA to MMRM, and do the nonparametric analysis of triglycerides for Part A as well.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
ADaM	Analysis Data Model
AE	Adverse event
AESI	Adverse event of special interest
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
AUDIT-C	Alcohol Use Disorders Identification Test-Concise
BAAT	Bile acid-CoA:amino acid N-acyltransferase
BARD	BMI-AST/ALT Ratio-Diabetes
BID	Twice daily
BMI	Body mass index
CAP	Controlled attenuation parameter
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence interval
CSR	Clinical Study Report
cT1	Iron-corrected T1
DBP	Diastolic blood pressure
DSMC	Data Safety Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic case report forms
ELF	Enhanced liver fibrosis
ET	Early termination
FIB-4	Fibrosis-4
GGT	Gamma-glutamyltransferase
HbA1c	Hemoglobin A1c
hs-CRP	High sensitivity C-reactive protein
HR	Heart rate
INR	International normalized ratio
IRT	Interactive Response Technology
ITT	Intent-to-treat
LDL	Low-density lipoprotein
LFC	Liver fat content
LLN	Lower limit of normal
LSM	Liver stiffness measurement
LSMean	Least square means
MAR	Missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified intent-to-treat
MMRM	Mixed-model repeated measures
MRI-PDFF	Magnetic resonance imaging-proton density fat fraction
NAFLD	Nonalcoholic fatty liver disease
NASH	Nonalcoholic steatohepatitis
PD	Pharmacodynamics

Abbreviation	Definition
PK	Pharmacokinetics
QD	Once daily
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SBP	Systolic blood pressure
SDTM	Study Data Tabulation Model
T2DM	Type 2 diabetes mellitus
TEAE	Treatment-emergent adverse event
TESAE	Treatment-emergent serious adverse event
TFLs	Tables, figures, and listings
ULN	Upper limit of normal
VAS	Visual analog scale
VLDL	Very low-density lipoprotein
WHO	World Health Organization

1 INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to provide a description of the statistical methods to be implemented for the analysis of data from the study with protocol number EYP001-202. The SAP will be finalized prior to the first database lock. Any deviations from the SAP after final database lock will be documented in the final Clinical Study Report (CSR).

This SAP covers both Part A and Part B of the study with a focus on what is needed for analysis of Part B. Refer to the SAP V1.0 for what was used to perform the analysis of Part A. Unless otherwise specified, the analyses of Part A and Part B will be performed separately.

2 STUDY OVERVIEW

2.1 Study Objectives and Endpoints

2.1.1 Primary Objective and Endpoint

Primary Objective	**Endpoint**
The primary objective is to determine the efficacy and safety profiles of EYP001a versus placebo on liver fat content (LFC) from baseline to Week 12 in patients with nonalcoholic steatohepatitis (NASH).	The primary efficacy variable is absolute change in LFC as measured by magnetic resonance imaging-proton density fat fraction (MRI-PDFF) from baseline to Week 12/Early Termination (ET). The secondary efficacy variables include the following: • Proportion of patients with ≥5% absolute reduction in LFC as measured by MRI-PDFF from baseline to Week 12/ET • Proportion of patients with ≥10% absolute reduction in LFC as measured by MRI-PDFF at Week 12/ET • Proportion of patients with ≥20% absolute reduction in LFC as measured by MRI-PDFF at Week 12/ET • Proportion of patients with a relative reduction in LFC ≥20% as measured by MRI-PDFF at Week 12/ET • Proportion of patients with a relative reduction in LFC ≥30% as measured by MRI-PDFF at Week 12/ET • Percent change (relative reduction) in LFC as measured by MRI-PDFF from baseline to Week 12/ET

2.1.2 Secondary Objectives and Endpoints

Secondary Objective	Endpoint
To explore the pharmacokinetics (PK) of EYP001a	EYP001a concentrations
To evaluate pharmacodynamic (PD) effects of EYP001a on bile acid-related markers as appropriate	Plasma levels of: • 7αhydroxy-4-cholesten-3-one; • fibroblast growth factor 19; • and total, primary, and secondary bile acids (such as o chenodeoxycholic acid, o deoxycholic acid, o lithocholic acid, o and/or others as appropriate).
To assess the effects of EYP001a on lipid and metabolic profiles	• Branched-chain amino acid concentrations • Hemoglobin A1c (HbA1c) • Homeostatic model assessment for insulin resistance (morning fasting glucose and insulin) • Lipoprotein insulin resistance index by nuclear magnetic resonance based on serum o cholesterol, o high-density lipoprotein cholesterol, o low-density lipoprotein (LDL) cholesterol, o triglycerides, o apolipoprotein B, o small dense-LDL, and o lipoprotein(a) • Lipoprotein particle size • Small LDL particle number • Small very low-density lipoprotein (VLDL) particle number • VLDL size
To assess the safety of EYP001a with statin coadministration	Safety and tolerability assessments will include the monitoring of

Secondary Objective	Endpoint
	• adverse events (AEs); • AEs of special interest (AESIs) (i.e., o pruritus [Part A only], o drug-induced liver injury [Part A only], o elevation of transaminases [Part B only], and o all muscle-related AEs); • serious AEs (SAEs); and findings from • physical examinations, • vital signs, • electrocardiograms (ECGs), and • clinical laboratory parameters.
To assess the effects of EYP001a on noninvasive biomarkers of liver fibrosis and inflammation	Change from baseline to Week 4, Week 8, and Week 14/ET, as applicable, in the following biomarkers of liver fibrosis and inflammation: • Alanine aminotransferase (ALT), • Aspartate aminotransferase (AST), • AST/ALT ratio, • adiponectin, • high sensitivity C-reactive protein (hs-CRP), • interleukin 6, • tumor necrosis factor alpha, • cytokeratin-18, • fibronectin, • hyaluronic acid, • procollagen type III N-terminal peptide, • tissue inhibitor of metalloproteinases-1 (and derived enhanced liver fibrosis [ELF] score), • Pro-C3, and • chitinase-3-like protein 1 (also known as YKL-40)

Additional secondary efficacy variables include the following:

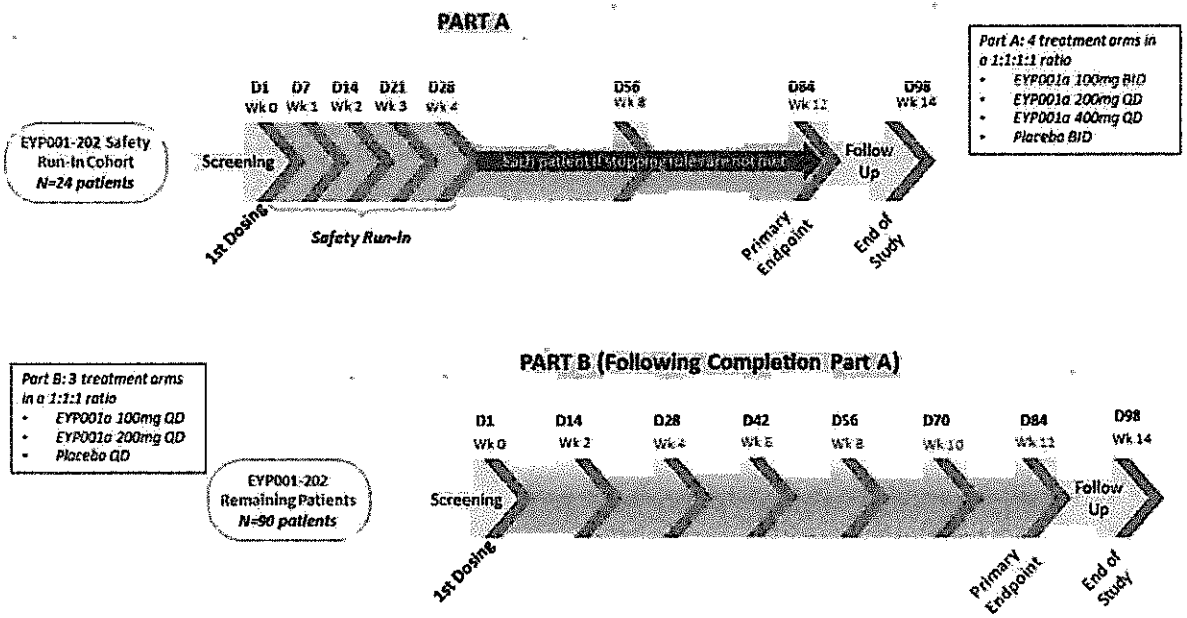
- Change and percent change in imaging-derived mean iron-corrected T1 (cT1) from baseline to Week 12
- Change and percent change in body mass index (BMI), waist circumference, and body weight
- Change and percent change in waist to hip ratio from baseline to Week 12 (Part B only)
- Change and percent change in waist to height ratio from baseline to Week 12 (Part B only)
- Derived non-invasive biomarkers of liver fibrosis and inflammation:
    - AST to platelet ratio index
    - Bile acid-CoA:amino acid N-acyltransferase (BAAT) Score
    - BMI-AST/ALT Ratio-Diabetes (BARD) Score
    - Nonalcoholic Fatty Liver Disease (NAFLD) Fibrosis Score
    - Fibrosis-4 (FIB-4) Score

2.2 Study Design

2.2.1 Overview

This is a 2-part, randomized, double-blind, multicenter, placebo-controlled study to evaluate the safety and efficacy of EYP001a in patients with NASH who likely have stage F2 to F3 fibrosis at approximately 50 global clinical sites. Overall, approximately 114 eligible patients will be enrolled: 24 patients in Part A (Safety Run-in Cohort), followed by 90 patients in Part B. In Part A, 24 patients will be randomized on Day 1 to 1 of 4 parallel treatment groups: 100 mg EYP001a twice daily (BID), 200 mg EYP001a once daily (QD), 400 mg EYP001a QD, or placebo BID. In Part B, 90 patients will be randomized on Day 1 to 1 of 3 parallel treatment groups: 100 mg EYP001a QD, 200 mg EYP001a QD, or placebo QD.

Randomization will be stratified by statin use (yes, no) and type 2 diabetes mellitus (T2DM) status (yes, no) at screening in Part A, and by statin use (yes, no) and LFC (stratification factor with 3 categories: LFC<16%, 16%≤LFC<22%, and 22%≤LFC) in Part B at screening.



BID = twice daily; D = Day; DSMC = Data Safety Monitoring Committee; QD = once daily; Wk = Week.

2.2.2 *Randomization and Blinding*

This study follows a randomized, double-blind, placebo-controlled design. Patients in Part A will be randomized via the Interactive Response Technology (IRT) system to 1 of 4 parallel treatment groups on Day 1 in a 1:1:1:1 ratio. However, prerandomization will occur within the IRT system at least 7 days prior to randomization. Patients in Part B will be randomized via the IRT system to 1 of 3 parallel treatment groups on Day 1 in a 1:1:1 ratio. Prerandomization is not applicable in Part B.

Both EYP001a and placebo will be provided as tablets for oral administration and will be identical in appearance.

Prerandomization/randomization information will be concealed from the Investigators and the patients until the end of the study, with the exception of an emergency situation involving a patient that required unblinding of the treatment assignment.

Stratification

Randomization will be stratified by statin use (yes, no) and T2DM status (yes, no) in Part A, and by statin use (yes, no) and LFC (stratification factor with 3 categories: LFC<16%, 16%≤LFC<22%, and 22%≤LFC) in Part B at screening. Randomization can occur after eligibility check on Day -3, at the earliest, and prior to dosing on Day 1 at the latest.

2.2.3 *Study Drug*

Formulation and Packaging

The drug product will be supplied as immediate-release coated tablets for oral administration. Excipients include lactose monohydrate, microcrystalline cellulose, crospovidone, sodium lauryl sulfate, silicon dioxide, and magnesium stearate. Placebo tablets will appear identical to EYP001a.

For Part A, all study drug (EYP001a and placebo) will be packaged in blister packs and wallets. One wallet will contain 2 weeks of study drug (i.e., 14 days of EYP001a or placebo) as per the randomization schedule. Morning and evening tablets will be clearly labeled.

For Part B, all study drug (EYP001a and placebo) will be packaged in blister packs and carton boxes. One box will contain 3 weeks of study drug (i.e., 3 blister strips or 21 days of EYP001a or placebo) as per the randomization schedule.

Study Drug Preparation and Dispensing

For Part A, study drug will be dispensed as 100 and 200 mg tablets for oral administration at Weeks 0, 2, 4, and 8 by blinded clinical site personnel. Patients will receive enough study drug to last until the next visit.

For Part B, study drug will be dispensed as 100 and 200 mg tablets for oral administration at Weeks 0, 2, 4, 6, 8, and 10 by blinded clinical site personnel. Patients will receive enough study drug to last until the next visit.

Study Drug Administration

EYP001a is only available for oral administration. The study drug should be administered with noncarbonated water.

For Part A, patients will take 2 tablets in the morning during breakfast and 1 tablet in the evening during dinner.

For Part B, patients will take 1 tablet in the morning during breakfast. On study visit days, patients should come fasted to the clinical site and take their morning dose during breakfast at the clinical site, following blood sample collection, except during intense PK/PD sampling on Days 1 and 14 of Part A, when dosing will occur with water only.

2.2.4 Sample Size Determination

The sample size of 114 randomized patients (n = 24 patients in Part A and n = 90 in Part B [n = 30 patients per treatment group]) is based on the efficacy endpoint: absolute change in LFC as measured by MRI-PDFF from baseline to Week 12, assuming that the treatment difference between each active treatment group with placebo is at least 5.1% with a common standard deviation of 5.8%.

With a 2-sample Z-test with the power of 0.8 and an overall family-wise 1-sided alpha of 0.025 (or each active versus placebo alpha = 0.0125 after Bonferroni adjustment), 26 patients per treatment group will be needed to complete the Week 12 Visit. With the consideration of an approximately 13% dropout rate before Week 12, 30 patients will be randomized per treatment group in Part B.

3 STATISTICAL METHODOLOGY

3.1 General Considerations

3.1.1 Treatment Groups

Data will be summarized and listed according to the randomized or actual treatment group, as appropriate.

In addition for Part A, data from the EYP001a treatment groups will be pooled and summarized. For analysis, due to the small number of patients by treatment group and as the PK results between group were found to be non-linear, with similar exposure for the EYP001a 200 mg QD and EYP001a 400 mg QD groups, only the pooled EYP001a treatment group will be compared with placebo.

In Part B, treatment groups will be summarized and analyzed separately. The EYP001a treatment groups will not be pooled.

3.1.2 Analysis Day

Analysis day will be calculated from the date of first dose of study drug. The day of the first dose of study drug will be Day 1, and the day immediately before Day 1 will be Day -1. There will be no Day 0.

Analysis Day will be calculated as:

Date of assessment – date of first dose of study drug

+ 1 (if date of assessment ≥ date of first dose of study drug)

3.1.3 Definition of Baseline

Baseline for the imaging parameters (LFC, cT1, liver stiffness measurement [LSM] and controlled attenuation parameter [CAP]) will be defined as the value obtained at the Screening Visit.

Baseline for ALT, AST, Total Bilirubin and alkaline phosphatase (ALP) will be defined as the average of all pre-dose values (i.e. those taken during the Screening Period and on Day 1 as assessed by the central laboratory). Only assessments taken within 12 weeks prior to Day 1 will be considered in the average calculation.

Baseline for the AST/ALT Ratio will be calculated using the derived baseline values for ALT and AST as described above.

Baseline for the other efficacy and safety variables will be defined as the last available value obtained up to the first dose of study drug.

For patients randomized and not treated, the baseline value will be the last available value on or before the day of randomization.

3.1.4 Analysis Visits

Refer to SAP V1.0 for Part A analysis visits.

<u>Part B</u>

For LFC, cT1, LSM and CAP, analysis visits will be assigned according to the labelled visit. I.e. "Baseline" analysis visit will be sourced from the "Screening" visit, and "Week 12/ET" analysis visit will be sourced from the "Week 12" or "ET" visit. Note that the ET assessment will be performed only for patients on study drug for ≥4 weeks.

For all other efficacy assessments with regular scheduled visits, analysis visits will be assigned according to the labelled visit. In addition to maximize the data available for assessments with regular scheduled visits, where possible, results from ET and unscheduled visits will be allocated to a post-baseline scheduled visit according to the following rules and visit windows

described in the table below (defined as the halfway point between the target visit days). If a valid value exists for the scheduled visit to which the ET/unscheduled visit has been allocated then the value from the scheduled visit will be used in the analysis, otherwise the values from the ET/unscheduled visits will be used in the analysis for the particular scheduled visit. 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 assessment from 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.

Analysis Visit and Target Analysis Day	Visit Windows (Analysis Days)		
	Group I	Group II	Group III
Baseline	See Section 3.1.3		
Day 14 (Week 2)	2 – 21	2 – 21	
Day 28 (Week 4)	22 – 35	22 – 42	2 – 42
Day 42 (Week 6)	36 – 49		
Day 56 (Week 8)	50 – 61	43 – 70	43 – 77
Day 70 (Week 10)	62 – 77		
Day 84 (Week 12)	78 – 91	71 – 91	
Day 98 (Week 14) EOS FUP	92+	92+	92+
Group I: ALT, AST, AST/ALT Ratio, ALP, Total Bilirubin, Gamma-glutamyltransferase (GGT) Group II: hs-CRP, CK, BMI, Waist Circumference, Body Weight, Waist to Height Ratio, Waist to Hip Ratio, and lipids and metabolic profile Group III: HbA1c and all other biomarkers of liver fibrosis and inflammation			

Note that parameters that are considered both efficacy and safety (ALT, AST, CK) will use analysis visit windowing for both the efficacy and safety summaries and analyses for Part A and Part B.

3.1.5 Absolute Change and Relative Change

Change (absolute) from baseline will be derived for each analysis visit as visit value minus baseline.

Percent (relative) change from baseline will be derived for each analysis visit as [(visit value minus baseline)/baseline] × 100.

3.1.6 Summary Statistics

Categorical data will generally be summarized with counts and percentages of patients. The denominator used for the percentage calculation will be clearly defined. Continuous data will generally be summarized with descriptive statistics including n (number of non-missing values), mean, standard deviation, median, and minimum, and maximum values

3.1.7 Hypothesis Testing

Unless stated otherwise, tests will be performed at the nominal alpha one-sided level of 0.025. 95% two-sided confidence intervals (CIs) will also be reported.

In Part A as only the pooled EYP001a treatment group is being compared to placebo, no adjustments for multiplicity are required.

In Part B, to adjust the 2 tests (each EYP001a dose and placebo) for multiplicity in order to control the family-wise type I error at 0.025 one-sided, the Bonferroni method will be used.

3.1.8 Evaluation of Site Effect

The number of subjects at each site is expected to be too small to evaluate site effect.

3.1.9 Handling of Dropouts and Missing Data

In general, few missing data are expected and, unless stated otherwise, they will not be replaced.

Missing values for the primary Week 12/ET endpoint, LFC, and the key secondary endpoint, cT1, will be imputed based on the missing at random (MAR) assumption using multiple imputation as the primary analysis. A sensitivity analysis for the primary endpoint, LFC, will also be carried out whereby the missing values will be imputed using the multiple imputation with tipping point method based on the missing not at random assumption. Further details of the multiple imputation methods will be provided in the relevant sections below.

3.2 Analysis Populations

Patients treated without or before being randomized will not be considered randomized and will not be included in any analysis populations

3.2.1 Randomized Population

The Randomized Population is defined as all patients who were randomized.

3.2.2 Intent-to-Treat (ITT) Population

The ITT Population includes all patients who are randomized and administered at least 1 dose of study drug.

3.2.3 Safety Population

The Safety Population includes all patients who are randomized and administered at least 1 dose of study drug.

3.2.4 Modified ITT Population

The Modified ITT (mITT) Population includes patients from the ITT Population who have valid baseline and Week 12/ET MRI-PDFF measurements of LFC.

3.2.5 Per-Protocol Population

The Per-Protocol Population includes all patients from the mITT Population without any major protocol violations. The analysis of LFC as measured by MRI-PDFF will be repeated on the Per-Protocol Population to test the robustness of results.

A major protocol violation is defined as any CSR reportable protocol deviation likely to affect the primary efficacy endpoint including the following categories (see Protocol Deviation Plan for further details):

- A patient that did not meet entry criteria;
- A patient that developed withdrawal criteria but were not withdrawn;

- A patient that received the wrong treatment or incorrect dose;
- A patient that received an excluded medication
- A patient that has less than 50% of morning compliance (part A) or less than 60% overall compliance (part B).

All patients with a CSR reportable protocol deviation will be reviewed and assessed for inclusion into the Per-Protocol Population; this review will be carried out by the Sponsor and Medical Monitor and agreed upon prior to the first database lock.

Excluded medications are detailed in the Protocol Section 5.6.1 Excluded Medications and/or Procedures (3rd and 4th bullet) that may have an effect on the primary efficacy endpoint:

- Vitamin E (>400 IU/day), glitazones, glucagon-like peptide-1 receptor agonists, ursodeoxycholic acid, or obeticholic acid within 90 days prior to screening
- Agents (including herbal over-the-counter weight loss preparations) or medications known to significantly impact body weight within 90 days prior to screening (e.g., sibutramine, phenetamine, and orlistat).

3.3 Subject Data and Study Conduct

3.3.1 Subject Disposition

This section describes patient disposition for both patient study status and the patient analysis populations. Note that all screen failures for the entire study will be considered together, i.e. combining screen failures from Part A and Part B.

A patient disposition summary will be provided for the Randomized Population and will include the following categories:

- Patients who were randomized
- Patients who received at least one dose of study drug
- Patients who completed the study
- Patients who withdrew from the study (and reasons for withdrawal)
- Patients whose primary reason for withdrawal was non-compliance due to COVID-19

Disposition events during the screening period will be summarized overall. Disposition events from the Day 1 (Week 0) onwards will be summarized by treatment group and overall.

Reasons for early withdrawal from the study (i.e. withdrawal after randomization) will be summarized by treatment group for the Randomized Population.

A disposition listing containing the randomization details (randomized treatment group and date of randomization), dates of first and last dose of study drug and completion status (with date and reasons for withdrawal for patients who withdrew) will be provided for all patients in the Randomized Population.

3.3.2 Protocol Deviations

Counts and percentages of patients with CSR reportable protocol deviations by deviation category will be summarized by treatment group for the Randomized Population. A similar summary will be provided for COVID-19 related CSR reportable protocol deviations.

A listing of CSR reportable deviations will also be produced for the Randomized Population.

Inclusion/Exclusion criteria deviations will also be listed for the Randomized Population.

3.3.3 Analysis Populations

Counts and percentages of patients in each analysis population will be summarized by treatment group and in total. Only counts will be displayed for the Randomized Population and the denominator for the other analysis populations will be based on the Randomized Population.

A listing of analysis population inclusion will also be provided.

3.3.4 Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized with descriptive statistics or counts and percentages of patients as appropriate by treatment group and in total for the ITT Population.

A similar summary will be provided for the mITT Population and Per-Protocol Population if the size of the population is different (>10%) from the size of the ITT Population for any treatment group.

Demographic characteristics include age, gender, ethnicity and race.

Baseline characteristics include:

- Stratification variables – statin use and T2DM status at screening (Part A), statin use and LFC category at screening (Part B);
- LFC as assessed from MRI-PDFF;
- cT1;
- LSM and CAP as assessed by FibroScan;
- Anthropometric parameters – height, weight, BMI and waist circumference, waist to hip ratio (Part B), waist to height ratio (Part B)
- Vital Signs – systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), Respiratory Rate and Temperature
- Alcohol use (Alcohol Use Disorders Identification Test-Concise [AUDIT-C] Total Score);
- Biomarkers of liver fibrosis and inflammation;
- Liver Function Tests – ALP, Total Bilirubin, GGT, International normalized ratio (INR) and Albumin.

Stratification variables will be sourced from the value used for randomization stratification, even if it is later found to be incorrect based on the patient data. Mis-stratification will be flagged in the listings.

3.3.5 Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 22. The number and percentage of patients with medical history will be summarized by

preferred term within system organ class and treatment group for the Randomized Population and presented by descending overall frequency of system organ class, and within system organ class by descending incidence of preferred term

All reported medical history will be listed.

3.3.6 Prior and Concomitant Medications

Prior and concomitant medications will be coded to World Health Organization (WHO) Drug Global B3, Mar-2019. A prior medication is defined as any medication taken by the patient before the first dose of study drug.

A concomitant medication is defined as any medication taken by the patient at the same time or after the first dose of study drug.

If a medication has incomplete start or stop dates, dates will be imputed to determine whether a medication should be considered concomitant or prior. If a medication start date is incomplete, the first day of the month will be imputed for missing day and January will be imputed for missing month. If a medication stop date is incomplete, the last day of the month will be imputed for missing day and December will be imputed for missing month. Incomplete start and stop dates will be listed as collected without imputation.

The number and percentage of patients taking prior and concomitant medications by preferred term within Anatomical Therapeutic Chemical (ATC) Class will be summarized by treatment group and in total based on the Safety Population.

The number and percentage of patients taking a prior or concomitant statin medication will also be summarized for the Safety Population and will include details of the type of statin taken.

All reported prior and concomitant medications, including statin medications, will be listed.

3.3.7 Study Drug Exposure and Compliance

Days of overall exposure to study drug will be summarized by treatment group for the Safety Population using descriptive statistics. Exposure, in days, will be calculated as:

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

Note that the exposure calculation is intended to describe the length of time a patient was exposed to study drug and therefore does not take study drug interruptions into account.

If the date of last dose of study drug is missing from the end of treatment electronic Case Report Forms (eCRF), for example if a subject is lost to follow-up, then the last dose of study drug will be imputed for the purpose of calculating exposure and compliance. The details of the imputation will be described in the programming specifications.

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

- ≤4 weeks (≤28 days)
- >4 to ≤8 weeks (29 - 56 days)
- >8 to ≤12 weeks (57 – 87 days)

Total percent compliance will be calculated as:

$$\frac{Number\ of\ tablets\ taken\ \times 100}{Expected\ number\ of\ tablets\ to\ be\ taken}$$

The number of tablets taken will be derived from the data recorded within the eCRF and will be calculated as the difference in the total number of tablets dispensed and the total number of tablets returned.

In Part A, only returned study drug wallets will be used for the calculation of compliance. In Part B if study drug is not returned, it will be imputed as 0 tablets returned for the purpose of calculating compliance. However if drug is dispensed on or after the date of last dose (such as in the case of a subject lost to follow-up and last dose is imputed), it may be excluded from the compliance calculation.

The expected number of tablets to be taken will be derived from the dates of first and last dose, assuming 3 tablets are taken per day (Part A) or assuming 1 tablet is taken per day (Part B).

The expected number of tablets to be taken will not take into account the duration of any temporary interruption to study treatment (e.g. due to AE).

In Part A, morning percent compliance will be calculated as:

$$\frac{Number\ of\ morning\ tablets\ taken\ \times 100}{Expected\ number\ of\ morning\ tablets\ to\ be\ taken}$$

In Part A, the number of morning tablets taken will derived from the data recorded within the eCRF as the difference in the total number of morning tablets dispensed and the total number of morning tablets returned. The total number of morning tablets dispensed is assumed to be two thirds of the total number of tablets dispensed.

In Part A, the expected number of tablets to be taken will be derived from the dates of first and last dose, assuming 2 tablets are taken in the morning per day.

Total and morning percent compliance with the study treatment will be summarized by treatment group for the Safety Population using descriptive statistics. Additionally, the number and percentage of patients within each treatment group with compliance in the following categories will be provided:

Total percent compliance

- < 80%
- 80 - 120% (inclusive)
- 120%

Morning percent compliance (Part A only)

- < 50%
- 50 - < 80%
- 80 - 120% (inclusive)
- 120%

All reported drug accountability, exposure, and compliance information will be listed.

3.4 Efficacy Assessment

In general, efficacy data will be summarized by treatment group. LFC as measured by MRI-PDFF and cT1 and all other efficacy data will be listed for the ITT Population.

3.4.1 Analyses of LFC as Measured by MRI-PDFF

Summary statistics for LFC at baseline, Week 12/ET and the change and percent change from baseline will be provided for the ITT, mITT and Per-Protocol Populations.

3.4.1.1 Primary Analysis

The primary efficacy outcome measure is the change in the LFC (as measured by MRI-PDFF) from baseline to Week 12/ET.

The primary efficacy analysis will be performed on the ITT Population. Missing MRI-PDFF values at Week 12/ET will be imputed based on the missing at random assumption using multiple imputation. Baseline MRI-PDFF will not be imputed; it is not expected to be missing for any patients.

The multiple imputation procedure and the analysis will be conducted in three separate steps: 1) Multiple Imputation, 2) ANCOVA Model based analysis by Imputed data set, and 3) combination of results using Rubin's rule.

1. Multiple imputation (using SAS PROC MI). Since baseline is required for inclusion and there is only one post-baseline MRI assessment, the missing pattern is monotone. The imputation model will include treatment, stratification factors, and baseline LFC. Note for Part B the baseline LFC is included, and so the LFC stratification factor will not be included in the imputation model. Missing data will be imputed 100 times to generate 100 complete datasets. Sample SAS code is below.

```
/*************************************************************
Part A:
      TRT01PN: 0=Placebo, 1=Pooled EYP001a
      STRAT1: T2DM (Y, N)
      STRAT2: Statin Use (Y, N)
Part B:
      TRT01PN: 0=Placebo, 1=EYP001a 100 mg QD, 2=EYP001a 200 mg QD
      STRAT1: Statin Use (Y, N)
      STRAT2: Not included
BASE: Baseline LFC
Week12: Week 12/ET LFC
*************************************************************/
proc mi data=ADEFF seed=345678 nimpute=100 out=midata1;
minimum=(. . . . 0) maximum=(. . . . 100);
class TRT01PN STRAT1 STRAT2;
monotone reg(WEEK12 = TRT01PN STRAT1 STRAT2 BASE /details);
var TRT01PN STRAT1 STRAT2 BASE WEEK12;
run;
```

2. Analysis of imputed datasets. Imputed datasets will be analyzed separately using the primary endpoint analysis model. The change in LFC from baseline to Week 12/ET will be analyzed using an analysis of covariance (ANCOVA) model with baseline LFC as a covariate and treatment and stratification groups as factors. Note for Part B the baseline

LFC is included as a covariate, and so the LFC stratification factor will not be included in the ANCOVA model. The least square means (LSMeans) for change from baseline to Week 12/ET LFC within the treatment groups and pairwise differences in LSMeans will be estimated along with their 95% CIs and p-values using Bonferroni method. Sample SAS code is below.

```
/*************************************************************
Part A:
      TRT01PN: 0=Placebo, 1=Pooled EYP001a
      STRAT1: T2DM (Y, N)
      STRAT2: Statin Use (Y, N)
Part B:
      TRT01PN: 0=Placebo, 1=EYP001a 100 mg QD, 2=EYP001a 200 mg QD
      STRAT1: Statin Use (Y, N)
      STRAT2: Not included
BASE: Baseline LFC
CHG: LFC change from Baseline to Week 12/ET
*************************************************************/
proc mixed data=miout;
by _Imputation_;
class TRT01PN STRAT1 STRAT2;
model CHG = TRT01PN STRAT1 STRAT2 BASE / ddfm=KR;
lsmeans TRT01PN / COV PDIFF CL;
run;
```

3. Combination of estimates across imputed datasets (using SAS PROC MINANALYZE). This will be performed using Rubin's rule and the overall treatment estimates vs. placebo along with their two-sided 95% CI, two-sided 97.5% CI, and p-value will be provided.

```
/*************************************************************
Part A: TRT01PN: 0=Placebo, 1=Pooled EYP001a
Part B: TRT01PN: 0=Placebo, 1=EYP001a 100 mg QD, 2=EYP001a 200 mg QD
*************************************************************/
proc mianalyze data=lsmeans;
by TRT01PN;
modeleffects estimate;
stderr stderr;
run;
```

The primary efficacy results will subsequently be analyzed by adding LSM and CAP values at enrollment as additional covariates.

Normality assumptions will be assessed via review of residual plots and non-parametric analyses may be explored if needed.

3.4.1.2 Sensitivity Analysis

Sensitivity analyses will be conducted by repeating the primary efficacy analysis based on:

- The ITT Population with missing MRI-PDFF values at Week 12/ET imputed using the multiple imputation with tipping point method based on the missing not at random assumption (further details provided below)
- The mITT Population

- The Per-Protocol Population

Tipping-point analysis

Robustness of the primary analysis results to the departure from the MAR assumption will be explored in the ITT population using the tipping-point analysis based on the pattern mixture model approach.

The considered pattern mixture model will introduce a penalty δ corresponding to the difference in mean change from baseline to Week 12/ET LFC between patients with missing data and patients with observed data.

Estimations will be performed using the same MI approach as described previously. In each treatment group, a penalty δ will be added to the imputed values (δ = 0 corresponds to the MAR assumption) after the PROC MI step. Sample SAS code is below.

```
/*************************************************************
Part A:
      TRT01PN: 0=Placebo, 1=Pooled EYP001a
      STRAT1: T2DM (Y, N)
      STRAT2: Statin Use (Y, N)
Part B:
      TRT01PN: 0=Placebo, 1=EYP001a 100 mg QD, 2=EYP001a 200 mg QD
      STRAT1: Statin Use (Y, N)
      STRAT2: Not included
&Shifti: δ, such as -1 to 1 by 0.1
&Shiftj: δ, such as -1 to 1 by 0.1
BASE: Baseline LFC
CHG: Change from Baseline to Week 12/ET LFC
*************************************************************/
proc mi data=ADEFF seed=345678 nimpute=100 out=midata1;
minimum=(. . . . -100) maximum=(. . . . 100);
class TRT01PN STRAT1 STRAT2;
monotone reg;
mnar adjust(CHG / shift=&Shiftj adjust=(TRT01PN='1'))
     adjust(CHG / shift=&Shifti adjust=(TRT01PN='0'));
var TRT01PN STRAT1 STRAT2 BASE CHG;
run;
```

To investigate how the conclusions depend on the adopted values δ, the testing will be repeated over a range of plausible values for the pairs (δ pooled EYP001a, δ Placebo for Part A; δ each EYP001a group, δ Placebo for Part B). The δ values range from -1 to 1 by 0.1, in each treatment group. The δ values may be adjusted depending on the actual data observed. Sample SAS code is below.

```
/*************************************************************
Part A:
      TRT01PN: 0=Placebo, 1=Pooled EYP001a
      STRAT1: T2DM (Y, N)
      STRAT2: Statin Use (Y, N)
Part B:
      TRT01PN: 0=Placebo, 1=EYP001a 100 mg QD, 2=EYP001a 200 mg QD
      STRAT1: Statin Use (Y, N)
      STRAT2: Not included
```

```
SHIFTEYP: δ, such as -1 to 1 by 0.1
SHIFTPBO: δ, such as -1 to 1 by 0.1
BASE: Baseline LFC
CHG: LFC change from Baseline to Week 12/ET
************************************************************/
proc GLM;
by _Imputation_ SHIFTEYP SHIFTPBO;
class TRT01PN STRAT1 STRAT2;
model CHG = TRT01PN STRAT1 STRAT2 BASE;
lsmeans TRT01PN / CL DIFF STDERR;
estimate "EYP001a vs Placebo" TRT01PN / 1 -1 / E;
run;

proc mianalyze data=diffs;
by SHIFTEYP SHIFTPBO;
modeleffects difference;
stderr stderr;
run;
```

Results will then be summarized using tables and a contour plot of the upper limit of 95% CI of LSMean difference between EYP001a and Placebo arms in change from baseline in LFC at Week 12/ET for the different δ will also be provided.

Then smallest value of δ EYP001a for a value of δ Placebo equal to 0 which will reverse the conclusion for the primary endpoint will be provided.

3.4.1.3 Key Secondary Analysis: Responders

A number of response criteria have been defined:

- ≥5% absolute reduction in LFC as measured by MRI-PDFF from baseline to Week 12/ET;
- ≥10% absolute reduction in LFC as measured by MRI-PDFF from baseline to Week 12/ET;
- ≥20% absolute reduction in LFC as measured by MRI-PDFF from baseline to Week 12/ET;
- Relative reduction in LFC ≥20% as measured by MRI-PDFF at Week 12/ET;
- Relative reduction in LFC ≥30% as measured by MRI-PDFF at Week 12/ET;

For each response criteria, the number and percentage of responders will be summarized by treatment group for the ITT Population; patients without a Week 12/ET MRI-PDFF assessment will be imputed as non-responders.

Each response criteria will be analyzed by a Cochran-Mantel-Haenszel test, stratified by statin use and T2DM status at screening for Part A, stratified by statin use and LFC stratification category for Part B. The responder analysis will be repeated using logistic regression models adjusting for treatment (pooled EYP001a and placebo for Part A, each EYP001a group and placebo for Part B), statin use and T2DM at screening for Part A, statin use stratification category for Part B, and baseline LFC. Note for Part B the baseline LFC is included as a covariate, and so the LFC stratification factor will not be included in the logistic regression model. The Cochran-Mantel-Haenszel test p-value will be reported. From the logistic regression, the odds ratio, two-sided 95% CI, and p-values will be reported.

3.4.1.4 *Key Secondary Analysis: Percent Change*

The main analysis of percent change from baseline will be performed on the ITT Population. As for the primary analysis, missing MRI-PDFF values at Week 12/ET will be imputed based on the missing at random assumption using multiple imputation.

The percent change from baseline to Week 12/ET in LFC will be analyzed for the ITT Population in similar ANCOVA model as described for the primary analysis. The analysis will be repeated as sensitivity analyses for the mITT and Per-Protocol Populations.

3.4.1.5 *Key Secondary Analysis: cT1*

The main analysis of change and percent change from baseline in cT1 will be performed on the ITT Population. As for the primary analysis, missing cT1 values at Week 12/ET will be imputed based on the missing at random assumption using multiple imputation. Missing baseline cT1 will not be imputed. Patients with missing baseline cT1 will not be included in the analysis.

The change and percent change from baseline to Week 12/ET in cT1 will be analyzed for the ITT Population using similar ANCOVA models to that described for the primary analysis, although baseline LFC covariate will be replaced with baseline cT1, and LFC stratification group (Part B) will be included as a factor. The analysis will be repeated as sensitivity analyses for the mITT and Per-Protocol Populations.

3.4.2 *Analysis of Other Efficacy Endpoints*

All other secondary efficacy endpoints will be summarized and analyzed based on the ITT Population. Descriptive statistics of the baseline, visit value and change from baseline to each post-baseline visit will be presented by treatment group for each secondary efficacy parameter. Percent change from baseline may be included where applicable.

3.4.2.1 *Biomarkers of Liver Fibrosis and Inflammation*

The change from baseline in the available biomarkers of liver fibrosis and inflammation at each analysis visit, as appropriate, will be analyzed using a mixed-model repeated measures (MMRM). The factors in the model will be treatment group (pooled EYP001a and placebo for Part A, each EYP001a group and placebo for Part B), stratification factors – statin use and T2DM status at screening for Part A or statin use and LFC stratification category for Part B, baseline value, visit and the treatment-by-visit interaction. An unstructured covariance matrix will be used (TYPE=UN). The LSMeans for change from baseline at each post-baseline visit will be estimated and compared between the treatment groups.

For Part B, non-parametric analysis will be used for hs-CRP. An unstratified Wilcoxon based analysis will be performed and the Hodges-Lehmann estimate along its 95% confidence intervals will also be provided.

The formulas for the derived noninvasive biomarkers will be documented in separate specifications. The derived noninvasive biomarkers, including ELF, will be summarized only. No formal statistical analysis will be conducted.

3.4.2.2 *Liver Stiffness Measurements (LSM, VCTE)*

The change and percent change from baseline to Week 12/ET in FibroScan results (LSM and VCTE) will be analyzed for the ITT Population using ANCOVA models with baseline LSM as a

covariate and treatment (pooled EYP001a and placebo) and stratification factors – statin use and T2DM status at screening as factors.

3.4.2.3 BMI, Waist Circumference, and Body Weight

The change and percent change from baseline in BMI, Waist Circumference and Body Weight at each analysis visit will be analyzed using a mixed-model repeated-measures (MMRM). The factors in the model will be treatment group (pooled EYP001a and placebo for Part A, each EYP001a group and placebo for Part B), stratification factors – statin use and T2DM status at screening for Part A or statin use and LFC stratification category for Part B, baseline value, visit and the treatment-by-visit interaction. An unstructured covariance matrix will be used (TYPE=UN). The LSMeans for change from baseline at each post-baseline visit will be estimated and compared between the treatment groups.

3.4.2.4 Waist to Hip Ratio and Waist to Height Ratio (Part B Only)

The change and percent change from baseline waist to hip ratio and waist to height ratio at each analysis visit will be analyzed using a mixed-model repeated-measures (MMRM). The factors in the model will be treatment group (each EYP001a group and placebo), stratification factors – statin use and LFC stratification category, baseline value, visit and the treatment-by-visit interaction. An unstructured covariance matrix will be used (TYPE=UN). The LSMeans for change from baseline at each post-baseline visit will be estimated and compared between the treatment groups.

3.4.2.5 Lipoprotein Profiles

Descriptive statistics of each safety laboratory parameter will be presented for baseline values and for values and the change and percent change from baseline at each post-baseline visit. Values outside the normal range will be categorized as H (high, i.e. above the normal range) or L (low, i.e. below the normal range) based on the central laboratory's normal reference range. The number and percentage of patients with values outside the normal range will be summarized by parameter for each post-baseline visit.

The change and percent change from baseline in each of Lipoprotein Profiling parameters (excluding triglycerides) at each analysis visit will be analyzed using a MMRM. The factors in the model will be treatment group (pooled EYP001a and placebo for Part A, each EYP001a group and placebo for Part B), stratification factors – statin use and T2DM status at screening for Part A or statin use and LFC stratification category for Part B, baseline value, visit and the treatment-by-visit interaction. An unstructured covariance matrix will be used (TYPE=UN). The LSMeans for change from baseline at each post-baseline visit will be estimated and compared between the treatment groups.

For analysis of the lipid and metabolic profile, only fasted values of cholesterol, lipoproteins, HDL-C, LDL-C and triglycerides will be included. The listing for these parameters will include the all results and an indicator of the fasting status of the patient at the time the sample was taken.

A sensitivity analysis will also be carried out on the Lipoprotein Profiling to investigate the effect of statin comedication; this analysis will use the same MMRM models as detailed above but with the addition of the interaction for statin use-by-treatment.

For Part A and B, non-parametric analysis will be used for triglycerides. An unstratified Wilcoxon based analysis will be performed and the Hodges-Lehmann estimate along its 95% confidence intervals will also be provided.

3.5 Pharmacokinetic Assessment

Further details concerning the display and analysis of the PK data will be included in a separate PK Analysis Plan and is out of the scope of this SAP.

3.6 Pharmacodynamic Assessment

Further details concerning the display and analysis of the PD data will be included in a separate PD Analysis Plan and is out of the scope of this SAP.

3.7 Safety Assessment

The safety data for this study include AEs, pruritus assessments, safety laboratory assessments, vital signs and anthropometrics, ECGs and physical examinations. The safety data will be summarized based on the Safety Population and presented by treatment group. Missing values will not be imputed. All reported safety data will be listed for the Randomized Population.

3.7.1 Adverse Events (AEs)

An AE is defined as any untoward medical occurrence in a clinical investigation patient administered a pharmaceutical product, which does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and/or unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of an investigational medicinal product, whether or not related to the investigational medicinal product.

For analysis, a treatment emergent AE (TEAE) is defined as a new or worsening AE after the first dose of study drug. Programmatically, this will be identified as any AE with a start date on or after the date of first dose. If part of the date for the onset/end of an AE is missing, but the existing parts allow determination of timing of AE onset relative to the first dose of study drug, then the AE will be classified as treatment-emergent or not per review of the existing parts of the date. If timing of the AE onset/end relative to the first dose of study drug cannot be determined, then the AE will be assumed to be treatment-emergent.

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

- Any AE
- Any TEAE
- Maximum severity of TEAEs
- Study drug-related TEAEs
- Maximum severity of study drug-related TEAEs
- TEAEs of special interest:
    - Part A: pruritus, Drug Induced Liver Injury (DILI) and muscle-related

- Part B: elevation of transaminases and muscle-related

- All serious AEs (SAEs)
- All serious TEAEs (TESAEs)
- All study drug-related TESAEs
- TEAEs leading to early discontinuation
- TEAEs leading to temporary discontinuation of study drug
- TEAEs leading to dose reduction

AEs will be coded using MedDRA v22. The number and percentage of patients with AEs, including the number of AEs, will be summarized by their MedDRA preferred term within system organ class and treatment group. For patient count summaries, multiple AE events with the same MedDRA coded terms (preferred term and system organ class) from the same patients will only be counted once. Similar summaries will be presented for study drug-related TEAEs and TESAEs.

The number and percentage of patients with TEAEs will be summarized by reported maximum severity within each MedDRA preferred term within system organ class and by treatment group.

A further summary table for TEAEs by preferred term and system organ class will be provided for patients in the Safety Population who were taking a statin; a separate summary table will be provided for patients in the Safety Population who were not taking a statin. Patients who were and were not taking a statin will be identified from the appropriate stratification variable.

Detailed listings will be provided for all AEs. Separate listings will also be provided for AEs of Special Interest, SAEs, TEAEs leading to early discontinuation, TEAEs leading to temporary discontinuation of study drug and TEAEs leading to dose reduction.

For both Part A and Part B, an additional summary of derived pruritus TEAEs will be provided for the following categories:

- Grade 1 (mild or localized), defined as (severity grade 1 and (flagged as pruritus AESI on the eCRF or preferred term="Pruritus generalised")) *or* preferred term = "Pruritus"
- Grade 2 (widespread and intermittent), defined as severity grade 2 *and* preferred term = "Pruritus generalised"
- Grade 3 (widespread and constant), defined as severity grade 3 *and* preferred term = "Pruritus generalised"

3.7.2 Pruritus Assessments

Part A

The pruritus visual analog scale (VAS) and 5-D (degree, duration, direction, disability and distribution) itch scale will be assessed at Days 1, 7, 14, 21, 28, 56, 84 and 98 and also at the ET Visit.

Descriptive statistics for the pruritus VAS and 5-D itch scale will be presented for baseline values and for values and the change from baseline at each post-baseline visit.

Part B

In Part B, pruritus will be reported as any other AE. If a patient reports pruritus, the visual analog scale and the 5-D questionnaire (degree, duration, direction, disability, and distribution) itch scale will be administered. Results will be listed.

3.7.3 Clinical Laboratory Tests

Serum chemistry, hematology, coagulation, lipid and metabolic profile parameters and urinalysis will be collected at the Screening Visit, Days 1, 7, 14, 21, 28, 56, 84, 98, and at the ET Visit (Part A) and at the Screening Visit, Days 1, 14, 28, 42 (ALT, AST, ALP, GGT, and total bilirubin only), 56, 70 (ALT, AST, ALP, GGT, and total bilirubin only), 84, 98, and at the ET Visit (Part B).

See Protocol Appendix B for analytes in the serum chemistry panel, additional chemistry parameters, hematology, coagulation, lipid and metabolic profile, and urinalysis.

The serum chemistry and additional chemistry parameters will be summarized and listed together.

Descriptive statistics of each safety laboratory parameter will be presented for baseline values and for values and the change from baseline at each post-baseline visit.

An additional summary table will be presented for CK over time for the patients in Safety Population who were taking a statin.

Values outside the normal range will be categorized as H (high, i.e. above the normal range) or L (low, i.e. below the normal range) based on the central laboratory's normal reference range. The number and percentage of patients with values outside the normal range will be summarized by parameter for each post-baseline visit.

Shift tables will also be presented for baseline vs. maximum post-baseline values and baseline vs. minimum post-baseline values for ALT, AST, Total Bilirubin and ALP. The shift tables for ALT and AST will include the following categories:

- < lower limit of normal (LLN) or Normal;
- > upper limit of normal (ULN) to ≤ 2 x ULN;
- > 2 x ULN ≤ 5 x ULN;
- > 5 x ULN to ≤ 8 x ULN;
- > 8 x ULN.

The shift tables for Total Bilirubin and ALP will include the following categories:

- < LLN or Normal;
- > ULN to ≤ 1.5 x ULN;
- > 1.5 x ULN to ≤ 2 x ULN;
- > 2 x ULN.

ALT, AST, and Total Bilirubin will be assessed using the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE). Shift tables will be presented for baseline vs. maximum post-baseline grades. The following categories are based on CTCAE Version 5.0 grading system:

Criteria	Grade 1	Grade 2	Grade 3	Grade 4
ALT increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
AST increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Total Bilirubin increased	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
Note: For purpose of deriving CTCAE grade, normal baseline is defined as a baseline value ≤1.0 x ULN, and abnormal baseline is defined as a baseline value >1.0 x ULN.				

An eDISH plot will also be provided for the Safety Population.

3.7.4 Vital Signs

Vital signs will be measured at the Screening Visit, Days 1, 7 (Part A only), 14, 21 (Part A only), 28, 56, 84 and 98 and at the ET Visit. Vital signs parameters include SBP, DBP, HR, respiratory rate and body temperature. At each visit, two measurements for blood pressure will be taken, the average of the two measurements will be used in the analysis.

Descriptive statistics of each vital signs parameter will be presented for baseline values and for values and the change from baseline at each post-baseline visit. These will be presented by treatment group for each parameter using the Safety Population.

3.7.5 Electrocardiograms

Single 12-lead ECGs will be performed at the Screening Visit, Days 7 (Part A only), 14, 21 (Part A only), 28, 56, 84 and 98 and at the ET Visit. ECG parameters include HR, PR interval, QRS duration, QT interval, QTcF interval and RR interval.

Descriptive statistics of each ECG parameter will be presented for baseline values and for values and the change from baseline at each post-baseline visit. Overall interpretation will also be summarized at each post-baseline visit. These will be presented by treatment group for each parameter using the Safety Population.

3.7.6 Physical Examinations

A complete physical examination (including general appearance; eyes, ears, nose and throat; head and neck; chest and lungs; cardiovascular; abdomen; musculoskeletal; lymphatic; dermatological; neurological systems; mental status; and extremities) will be carried out at the Screening Visit. Brief, symptom directed physical examinations will be carried out at Days 1, 7 (Part A only), 14, 21 (Part A only), 28, 56, 84 and 98 and at the ET Visit.

A complete listing of the physical examination results will be provided for the Safety Population.

4 ANALYSIS TIMING

4.1 Interim Analysis

Two interim analyses were conducted during the course of the study:

- The first interim analysis was to be performed at the end of the Safety Run-in Period, after the last of the 24 patients randomized in Part A completed Day 28 (Week 4). It was anticipated that at the time of the first interim analysis, approximately 10 patients would have completed Day 84 (Week 12) of Part A. This analysis was completed in November, 2019, when 19 subjects had completed Week 4 and 11 subjects had completed Week 12.
- A second interim analysis was performed when 50% of patients in Part B completed Day 56 (Week 8). This analysis was completed in February, 2021.

Refer to the Interim Analysis Plan for details.

4.2 Pre-final Analysis

Part A: After the database is frozen and exclusions from analysis populations have been finalized, the Part A randomized treatment assignments will be unblinded and the Part A pre-final analysis will be generated.

Part B: Database lock and analysis will occur in two steps as described below. The purpose of conducting the database lock and analysis in two steps is to obtain the primary analysis results of the 12-week treatment period as soon as possible for corporate goals. The trigger for the first step database lock is when the last patient finishes Week 12. Patients will no longer be on treatment for the remainder of the study follow-up period, therefore it is not considered to be a substantial risk to the patients or study data to conduct the primary analysis prior to the final database lock.

1. First step: efficacy analyses up to Week 12 - This analysis will be conducted on all randomized patients when all patients will have all their MRI-PDFF, cT1, and laboratory data up to the Week 12 analysis window collected and validated (i.e. SDV'd, reconciled, and locked). The efficacy analyses will be performed up to the Week 12 time point. In other words, the Analysis Data Model (ADaM) efficacy datasets will only include data through Week 12.

   Analyses of MRI-PDFF and cT1 endpoints up to the Week 12 time point will correspond to the final analyses for these endpoints. Any other analyses using MMRM, such as laboratory data, up to the Week 12 time point will not exactly correspond to the final analyses for these endpoints. This is because the analysis model is MMRM and includes

Week 14 time points in the model, which will not be included for the first step database lock. After the final database lock when all Week 14 data is available and added to the MMRM, results for all visits used in the MMRM may change.

The efficacy data to be analyzed up to Week 12 includes MRI-PDFF, cT1, LSM, ALT,LDL-C, and body weight.

The safety analyses will be performed on all safety data collected up to the common cut-off date. For this analysis, the common cut-off date is defined as the date of the first database lock that occurs after the last Week 12 visit. The safety analysis will be updated after the second step database lock.

The safety data to be analyzed includes adverse events and GGT.

Summaries of demographics and disposition will also be provided.

ENYO Pharma and Medpace biostatistics team (including Study Data Tabulation Model [SDTM] programmers) will be unblinded at the time of the first step database lock in order to perform the analysis. The analysis will be conducted in a secure location with limited access in order to maintain the blind for the remainder of the study team.

2. Second step: final analysis - This analysis will be conducted at the end of the study and will consist in the final analysis of efficacy data at time points beyond the Week 12 time point and final safety analysis.

Two database locks will be done:

- First database lock (for first step analysis): will include all available data on all randomized patients up to the common cut-off date as defined above. All data needed for determining the Per-Protocol Population and for the first step analysis described above will be included in the first database lock.
- Final database lock (for second step analysis): will include all data, including follow-up, for all screened patients.

Topline TFLs will be provided approximately 1 week after first database lock and unblinding and the full set of pre-final TFLs will be provided approximately 3 weeks after final database lock.

In addition to TFLs, SDTM data and ADaM data along with associated files will be provided as requested ENYO. The SDTM database includes all subjects from Part A and Part B. The ADaM datasets may separate Part A and Part B subjects.

Associated files may include: annotated CRFs, SDTM specifications, SDTM programs, ADaM specifications, ADaM programs, TFL programs, and Clinical Data Interchange Standards Consortium (CDISC) Define packages for both SDTM and ADaM data.

4.3 Final Analysis

After all comments on the pre-final analysis have been resolved and the study database is declared final, the final analysis will be generated. Final tables, figures, and listings (TFLs) will be provided approximately 1 week after the study database is declared final. This includes the full set of Part A TFLs and Part B TFLs. If there were no changes to the pre-final analysis or the study database, the pre-final TFLs may be considered final.

5 CHANGES FROM PROTOCOL-SPECIFIED STATISTICAL ANALYSES

For Part B analysis of LFC as measured by MRI-PDFF, the ANCOVA model specified in the Protocol includes both baseline LFC as a covariate and LFC stratification (3 categories) as a factor. Since this is redundant information, LFC stratification as a factor will be dropped as a covariate from the model; baseline LFC as a covariate is sufficient.

After Part A analysis was completed, the Per-Protocol Population definition was updated to require subjects to finish the 12-week Treatment Period. However, this part of the definition is excluded for analysis. Instead, the Part B Per-Protocol Population is clarified to require ≥60% study drug compliance.

There are some other minor changes from the protocol-specified statistical analyses detailed in this SAP:

- Week 12 included for biomarkers of liver fibrosis and inflammation, as appropriate;
- Study endpoints and definition of mITT Population updated to clearly allow patients with ET assessments (not just Week 12 assessments) to be included.
- HbA1c is not analyzed at Week 2 and Week 12 with the other lipid and metabolic profile parameters.

6 PROGRAMMING SPECIFICATIONS

Analyses will be performed using SAS® version 9.4 or higher. All available data will be presented in patient data listings which will be sorted by patients and visit date as applicable. Detailed Programming Specifications will be provided in a separate document(s).