# Statistical Analysis Plan for Protocol 747-303

## A Phase 3, Double-Blind, Randomized, Long-Term, Placebo-Controlled, Multicenter Study Evaluating the Safety and Efficacy of Obeticholic Acid in Subjects with Nonalcoholic Steatohepatitis

## The REGENERATE Study

## RandomizEd Global Phase 3 Study to Evaluate the Impact on NASH with FibRosis of Obeticholic Acid TreatmEnt

## OBETICHOLIC ACID (OCA)

| | |
|---|---|
| **Protocol Version and Date:** | 11.0: 04 May 2021 |
| **Phase:** | Phase 3 |
| **Methodology:** | Double-Blind, Randomized, Placebo-Controlled Study |
| **Sponsor:** | Intercept Pharmaceuticals, Inc.<br>305 Madison Avenue<br>Morristown, NJ 07960<br>Tel: 646-747-1000<br>Fax: 858-558-5961 |
| **Sponsor Representative:** | |
| **Analysis Plan Date:** | 31 January 2023 |
| **Analysis Plan Version:** | 2.0 |

NCT Number: NCT02548351

This NCT number has been applied to the document for purposes of posting on Clinicaltrials.gov

# APPROVALS:

**Authored by:**

[redacted], MD, MPH, PhD
[redacted], Biostatistics
Intercept Pharmaceuticals, Inc.

Date

[redacted], MS
[redacted] Biostatistician
Intercept Pharmaceuticals, Inc.

Date

**Approved by:**

[redacted], PhD
[redacted], Biostatistics & Data Management
Intercept Pharmaceuticals, Inc.

Date

[redacted], MD
[redacted], Clinical Development
Intercept Pharmaceuticals, Inc.

Date

[redacted], MD
[redacted], Regulatory Affairs
Intercept Pharmaceuticals, Inc.

Date

# TABLE OF CONTENTS

## LIST OF TABLES



## LIST OF FIGURES

# LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

| Abbreviation | Term |
|---|---|
| AE | adverse event |
| AESI | adverse events of special interest |
| AKI | acute kidney injury |
| ALP | alkaline phosphatase |
| ALT | alanine aminotransferase |
| ANCOVA | analysis of covariance |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| ATC | Anatomical Therapeutic Chemical |
| [redacted] | [redacted] |
| AUDIT | Alcohol Use Disorders Identification Test |
| BLQ | below limit of quantitation |
| BMI | body mass index |
| CEC | clinical events committee |
| CI | confidence interval |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| CMH | Cochran-Mantel-Haenszel |
| CRL | Complete Response Letter |
| CRP | C-reactive protein |
| CRN | Clinical Research Network |
| CS | clinically significant |
| CSR | clinical study report |
| CV | cardiovascular |
| CV% | coefficient of variation |
| DAP | Data Access Plan |
| DCO | Data cutoff Date |
| DILI | drug induced liver injury |
| DMC | Data Monitoring Committee |
| DOB | date of birth |
| DOIC | date of informed consent |
| eCRF | electronic case report form |

| Abbreviation | Term |
|---|---|
| EAIR | Exposure adjusted incidence rate |
| ECG | electrocardiogram |
| EDC | Electronic Data Capture |
| eDISH | Evaluation of drug-induced serious hepatotoxicity |
| [redacted] | [redacted] |
| FDA | Food and Drug Administration |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| FRS | Framingham Risk Score |
| GGT | gamma-glutamyl transferase |
| glyco-OCA | glycine 6α-ethyl chenodeoxycholic acid |
| [redacted] | [redacted] |
| HDL | high density lipoprotein |
| IA | Interim analysis |
| ICH | International Conference on Harmonization |
| IE | Intercurrent Event |
| in | inches |
| IgM | immunoglobulin M |
| IIEF-15 | International Index of Erectile Function-15 |
| INR | international normalized ratio |
| | |
| IQR | interquartile range |
| ITT | Intent-to-Treat |
| IWQoL-Lite | Impact of Weight on Quality of Life-Lite |
| IWRS | interactive web response system |
| kg | kilogram |
| KM | Kaplan-Meier |
| lb | pounds |
| LDL | low density lipoprotein |
| LLN | lower limit of normal |
| LS | least-square |
| MACE | major adverse cardiovascular events |
| MELD | Model for End Stage Liver Disease |
| MedDRA | Medical Dictionary for Regulatory Activities |

| Abbreviation | Term |
|---|---|
| mITT | modified Intent-to-Treat |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| µmol | micromole |
| mg | milligram |
| MMRM | mixed-effect repeated measures model |
| MW | molecular weight |
| NAFLD | nonalcoholic fatty liver disease |
| NAS | nonalcoholic fatty liver disease activity score |
| NASH | nonalcoholic steatohepatitis |
| NCS | not clinically significant |
| NDA | New Drug Application |
| NEC | Not Elsewhere Classified |
| NFS | nonalcoholic fatty liver disease fibrosis score |
| ng | nanogram |
| OCA | obeticholic acid |
| PBC | primary biliary cirrhosis |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| PT | preferred term |
| QTcF | QT interval corrected by the Fridericia's formula |
| RR | time between 2 consecutive R waves |
| SAE | serious adverse event |
| [redacted] | [redacted] |
| SD | standard deviation |
| SEM | standard error of the mean |
| SI | standard international unit |
| SMQ | Standardized MedDRA Query |
| SOC | system organ classes |
| [redacted] | [redacted] |
| TEAE | treatment-emergent adverse event |
| [redacted] | [redacted] |
| TNF-α | tumor necrosis factor-alpha |

| Abbreviation | Term |
|---|---|
| TZD | thiazolidinedione |
| U/L | units per liter |
| ULN | upper limit of normal |
| WHO | World Health Organization |

# 1. INTRODUCTION

This document outlines the statistical methods to be implemented for the Month 18 Interim Analysis and End of Study Analysis of Study 747-303 (Phase 3 data collected within the scope of the Intercept Pharmaceuticals, Inc. [the Sponsor], sponsored protocol). Access to unblinded analyses is specified in the data access plan (DAP) and Unblinding Plan (24 Mar 2022). Any deviations from the methods specified in this Statistical Analysis Plan (SAP) will be documented in the clinical study report (CSR).

On 26 Jun 2020, the Food and Drug Administration (FDA) issued a Complete Response Letter (CRL) to New Drug Application (NDA) 212833. Resolution of deficiencies outlined in the CRL required additional histological and safety analyses, which are outlined in SAP Addendum V2.0 (Dated 25 Apr 2022).

It is important to note that all analyses following the CRL that were used to support a re-submission are described in the SAP addendum dated 25 Apr 2022. The Month 18 Interim Analyses that refer to sections in this document (Section 2.1.1, Section 2.3.1, Section 3.4.1) were done for the original NDA submission. The updates in this SAP document are relevant to End of Study analyses.

Key changes in this document compared to SAP V1.0 (dated 24 Jan 2019) are summarized in Section 11.

# 2. INFORMATION FROM THE STUDY PROTOCOL

## 2.1. Study Objectives

This section provides the study objectives at the Month 18 Interim Analysis and at the End of Study Analysis.

### 2.1.1. Month 18 Interim Analysis

#### 2.1.1.1. Primary Objective Assessed at 18 Months

To evaluate the effect of obeticholic acid (OCA) compared to placebo on histological improvement of nonalcoholic steatohepatitis (NASH) by assessing the following endpoints using NASH Clinical Research Network (CRN) scoring criteria:

- Improvement of fibrosis by at least 1 stage with no worsening of NASH (no worsening of hepatocellular ballooning, no worsening of lobular inflammation and no worsening of steatosis)
- Resolution of NASH with no worsening of fibrosis

NASH resolution is defined as the overall histopathologic interpretation of 1) “no fatty liver disease” or 2) “fatty liver disease (simple or isolated steatosis) without steatohepatitis” AND a nonalcoholic fatty liver disease (NAFLD) activity score (NAS) of 0 for ballooning and 0-1 for inflammation.

#### 2.1.1.2. Secondary Objectives Assessed at 18 Months

- To evaluate the effect of OCA compared to placebo on histological improvement of NASH by assessing the following using NASH CRN scoring criteria:
  - Improvement of fibrosis by at least 1 stage AND/OR resolution of NASH, without worsening of either
  - No worsening of fibrosis AND no worsening of NASH
  - Progression to cirrhosis
  - Improvement of fibrosis by at least 2 stages
  - Improvement of each key histological feature of NASH by at least 1 point (steatosis, lobular inflammation, and hepatocellular ballooning)
  - Improvement of NAS by at least 2 points with no worsening of fibrosis
  - Improvement of fibrosis and resolution of NASH as a composite endpoint and as defined by both endpoints being met in the same subject
  - Resolution of fibrosis
- To evaluate the effect of OCA compared to placebo on liver biochemistry and markers of liver function

- Incidence of adjudicated cardiovascular events
- Safety and tolerability (treatment-emergent adverse events (TEAEs), electrocardiograms (ECG), vital signs, clinical laboratory assessments)

### 2.1.2. End of Study Analysis:

#### 2.1.2.1. Primary Objective Assessed at End of Study

- To evaluate the effect of OCA compared to placebo on all-cause mortality and liver-related clinical outcomes as measured by the time from randomization to the first occurrence of any of the following adjudicated events (clinical outcomes composite endpoint):
  - Death (all cause)
  - Model of end stage liver disease (MELD) score ≥15
  - Liver transplant
  - Hospitalization (as defined by a stay of ≥24 hours) for onset of:
    - Variceal bleed
    - Hepatic encephalopathy (as defined by a West Haven score of ≥2)
    - Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis)
  - Ascites secondary to cirrhosis and requiring medical intervention (eg, diuretics or paracentesis)
  - Histological progression to cirrhosis

#### 2.1.2.2. Secondary Objectives Assessed at End of Study

- To evaluate the effect of OCA compared to placebo on histological improvement of NASH by assessing the following endpoints using NASH CRN scoring criteria:
  - Improvement of fibrosis by at least 1 stage with no worsening of NASH

  - Resolution of NASH with no worsening of fibrosis
  - Improvement of fibrosis by at least 1 stage AND/OR resolution of NASH, without worsening of either
  - No worsening of fibrosis AND no worsening of NASH
  - Improvement of fibrosis by at least 2 stages
  - Improvement of each key histological feature of NASH by at least 1 point (steatosis, lobular inflammation, and hepatocellular ballooning)
  - Improvement of NAS by at least 2 points with no worsening of fibrosis
  - Improvement of fibrosis and resolution of NASH as a composite endpoint and as defined by both endpoints being met in the same subject
  - Resolution of fibrosis
- To evaluate the effect of OCA compared to placebo on liver biochemistry and markers of liver function

#### 2.1.2.4. Safety Objectives Assessed at End of Study

- To evaluate the effect of OCA compared to placebo on:
  - Incidence, exposure adjusted incidence rate and time to treatment-emergent pruritus
  - Incidence, exposure adjusted incidence and time to adjudicated cardiovascular events
  - Incidence and exposure adjusted incidence adjudicated acute kidney injury (AKI)

- Incidence, exposure adjusted incidence and time to adjudicated events of hepatic injury
- Long-term safety and tolerability (TEAEs, AESIs, ECGs, vital signs, clinical laboratory assessments)

## 2.2. Study Design

### 2.2.1. Overall Study Design

This Phase 3, double-blind, randomized, long-term, placebo-controlled, multicenter international study will evaluate the effect of OCA on histological improvements of NASH, all-cause mortality, and liver-related clinical outcomes. According to the original design, the study will enroll approximately 2370 subjects with NASH. The population will comprise approximately 2085 subjects with biopsy-confirmed, precirrhotic NASH and evidence of liver fibrosis, including approximately 60% with fibrosis stage 3 and 40% with fibrosis stage 2. An additional cohort of approximately 285 subjects with fibrosis stage 1, and at least 1 accompanying comorbidity, will also be enrolled to gather information on the safety of OCA and progression of liver disease in this population as detailed in this document. A planned interim analysis (Month 18 Interim Analysis) will be performed after a minimum of 750 randomized subjects (the first sequential) with fibrosis stage 2 or stage 3 would have reached their actual/planned Month 18 Visit (including subjects who discontinued before reaching the planned Month 18 Visit).

Following the Month 18 Interim Analysis, the study will be continued in a blinded fashion, and subjects will be followed over an extended period for clinical outcomes to confirm clinical benefit as part of the End of Study Analysis. The End of Study Analysis will evaluate the effect of OCA on mortality and liver-related clinical outcomes, as well as the long-term safety of OCA as described in the protocol after accrual of approximately 288 adjudicated clinical outcome composite events combined in the OCA 25 mg and placebo groups for subjects with fibrosis stage 2 or stage 3 based on the group sequential design with one Interim Analysis at 66% of information fraction. The sample size required for the updated analysis can be found in Section 3.1.

Subjects will be screened for a period of up to 12 weeks before entering the study. Approximately 2480 subjects with NASH who meet all eligibility criteria will be randomized to receive OCA 10 mg, OCA 25 mg, or matching placebo in a 1:1:1 ratio for the duration of the study, in conjunction with local standard of care. Randomization of subjects with fibrosis stage 2 or stage 3 will be stratified by presence of diabetes at enrollment [yes/no] and use of thiazolidinediones (TZDs) or vitamin E at baseline [yes/no]. Biopsy-confirmed NASH will be determined by central reading of liver histology at Screening, and all subsequent histology assessments will be performed centrally in a blinded fashion. At Month 18 or early termination, biopsy slides will be paired-read with the screening biopsy in a blinded fashion to minimize temporal bias of the Month 18 biopsy read.

Following Month 18 visits, clinical visits will occur every 6 months for the remaining duration of the study. For any subject who develops an AE of hepatic injury or decompensation, or if cirrhosis is suspected based on criteria other than liver biopsy or clinical outcomes, a liver biopsy is recommended and will be read using Consensus Method except that in a case where a biopsy

was already obtained within 6 months of the suspected liver injury or decompensation event. All deaths and potential liver-related clinical outcomes including clinical events and histological findings of potential disease progression to cirrhosis adjudicated by independent Hepatic Outcome Committee will be used for the End of Study analysis.

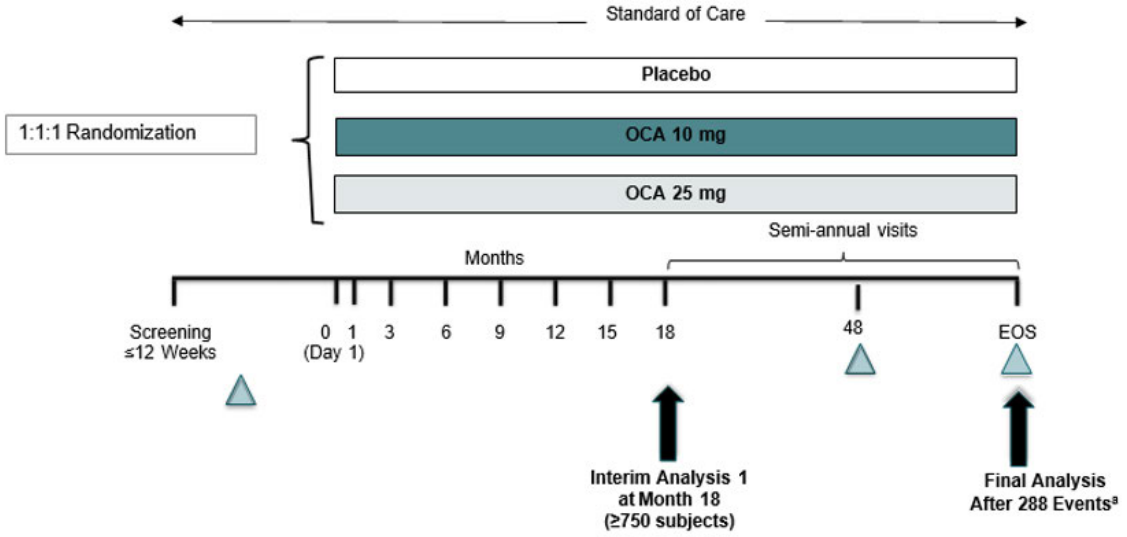
Investigational product (ie, OCA or placebo) will be taken orally, with water, once daily. Subject safety and laboratory assessments will be evaluated at clinical visits as detailed in the protocol.

With the exception of liver transplant, if a subject experiences a suspected or confirmed clinical outcome event, he or she should continue the regular visit schedule and continue taking investigational product as long as none of the criteria for discontinuation outlined in the protocol have been met. Subjects who discontinue investigational product due to Investigator or Sponsor decision are expected to be followed through to study closure (or at the discretion of the Sponsor).

Details of study procedures and assessments are described in the protocol.

### 2.2.2. Study Design Diagram

**Figure 1: Study Design Schematic**




EOS = End of study; OCA=obeticholic acid
▲ Biopsy (Subjects without a liver biopsy performed within 6 months before Day 1 will have a biopsy at the second Screening Visit).
* Number of hepatic clinical outcomes adjudicated events accrued in placebo and OCA 25 mg groups combined with one Interim Analysis at 66% of information fraction.

### 2.2.3. Randomization Methodology

This study will be conducted in a double-blind, placebo-controlled manner. Subjects will be randomized 1:1:1 to placebo, OCA 10 mg, or OCA 25 mg. Randomization will be stratified by presence of diabetes at enrollment [yes/no] and use of TZDs or vitamin E at baseline [yes/no].

The randomization will be based on a predefined randomization code (generated by the Sponsor or designee) using an interactive web response system (IWRS), which the Sponsor intends will serve as a computer-based subject registration system at Screening and Day 1. The IWRS will serve as an investigational product inventory and management system and may be integrated with the study database.

The Investigator or designee will be required to register the subject in the IWRS and may be prompted to provide subject data necessary to properly randomize and allocate the subject to treatment. A randomization number will be assigned and investigational product dispensing information (ie, bottle number[s]) will be provided.

#### 2.2.4. Monitoring Guidelines and Unblinding

Once enrolled, subjects will be monitored closely by reviewing their AEs, signs and symptoms, and laboratory values especially for potential hepatic injury and/or decompensation, pruritus, gallstone disease, symptomatic cholelithiasis and/or cholecystitis or pancreatitis, renal impairment and nephrolithiasis. Dosages for investigational product should be maintained consistent with the protocol during for the duration of the study. Any dose adjustments, interruptions, discontinuation, and any rechallenge of investigational product should be implemented per the criteria outlined in the protocol based on safety or tolerability concerns (see Appendix D).

Subjects that discontinue investigational product are expected to continue in the study until study termination, and the study will only terminate at the time when the required number of adjudicated events has accrued (Section 3).

Subjects who discontinue investigational product are expected to continue in the study and complete all scheduled study visits and assessments and are to be followed for the occurrence of outcome (hepatic) events.

The study will be conducted in a blinded manner with the exception a selected group of individuals who are unblinded for the Month 18 Interim Analysis as described in the DAP and Unblinding Plan. After the End of Study database lock, the study data will be completely unblinded.

## 2.3. Study Endpoints

### 2.3.1. Efficacy Analysis at Month 18 Interim Analysis

#### 2.3.1.1. Primary Efficacy Endpoints

The primary efficacy analysis for the Month 18 Interim Analysis will compare subjects treated with OCA (10 mg and 25 mg) versus placebo on histological improvement of nonalcoholic steatohepatitis (NASH) by assessing both of the following using NASH Clinical Research Network (CRN) scoring criteria:

- The percentage of subjects with improvement of fibrosis by 1 stage or more, no worsening of hepatocellular ballooning, no worsening of lobular inflammation, and no worsening of steatosis

- The percentage of subjects with resolution of NASH and no worsening of fibrosis

The study will meet its primary objective if either of the above primary endpoints is met.

#### 2.3.1.2. Secondary Efficacy Endpoints

##### 2.3.1.2.1. Histological Analysis:

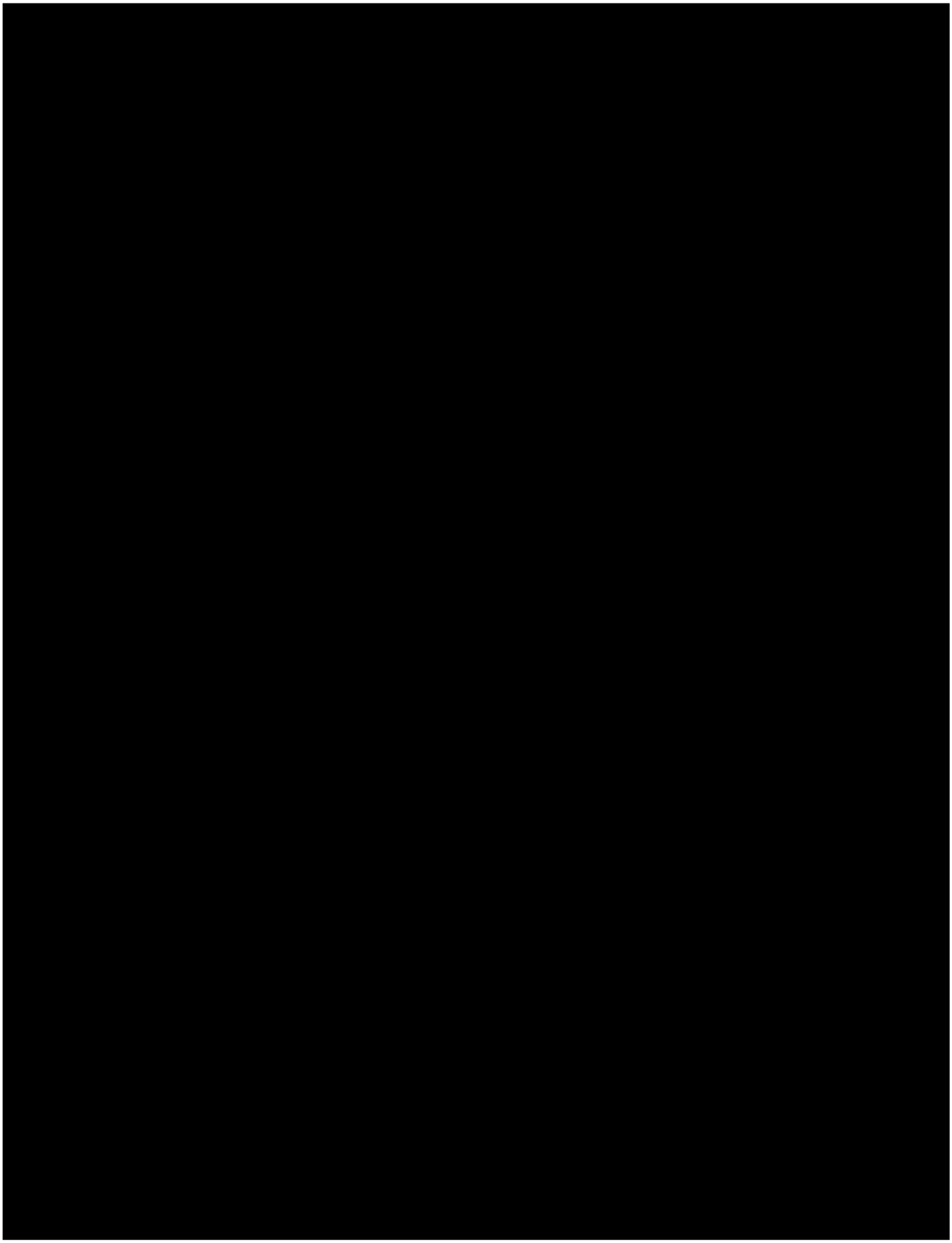
The analysis of the secondary histological efficacy endpoints for the Month 18 Interim Analysis, including the key secondary endpoint will compare subjects treated with OCA (10 mg and 25 mg) versus placebo on histological improvement by assessing the following using NASH CRN scoring criteria:

Key Secondary Endpoint:

- Percentage of subjects with improvement of fibrosis by at least 1 stage AND/OR resolution of NASH, without worsening of either

Secondary Endpoints:

- Percentage of subjects with no worsening of fibrosis AND no worsening of NASH
- Percentage of subjects with progression to cirrhosis
- Percentage of subjects with improvement of fibrosis by at least 2 stages
- Percentage of subjects with improvement of each key histological feature of NASH by at least 1 point (steatosis, lobular inflammation, and hepatocellular ballooning)
- Percentage of subjects with improvement of NAS by at least 2 points with no worsening of fibrosis
- Percentage of subjects with improvement of fibrosis and resolution of NASH as a composite endpoint and as defined by both endpoints being met in the same subject
- Percentage of subjects with resolution of fibrosis



### 2.3.2. Efficacy Analysis at End of Study

#### 2.3.2.1. End of Study Primary Estimands and Efficacy Endpoints

The primary estimands at the end of study (EOS) will be the difference in time from randomization to the first occurrence of one of the post-randomization events listed below (clinical outcome composite endpoints). This will be assessed between OCA and placebo in adult subjects with (NASH) at fibrosis stage 2 or 3.

- Death (all-cause)
- MELD score ≥15
- Liver transplant
- Hospitalization (as defined by a stay of 24 hours or greater) for onset of:
  - Variceal bleed
  - Hepatic encephalopathy (as defined by a West Haven score of ≥2)

  - Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis)
- Ascites secondary to cirrhosis and requiring medical intervention (eg, diuretics or paracentesis)
- Histological progression to cirrhosis

For subjects who had an intercurrent event like investigational product interruption or dose adjustment or investigational product discontinuation and continued in the study, the events after those intercurrent events will still be evaluated for the analysis following Treatment Policy. If an intercurrent event like study discontinuation due to various reasons, withdrawal from study, or lost-to-follow-up occurred, the last available assessment prior to the intercurrent events will be used for the analysis following the While on Treatment Policy.

**Table 1: Estimands for Clinical Composite Outcome**

| Estimands | Population | Endpoint | Intercurrent Events | Strategy for IEs | Population-Level Summary |
|---|---|---|---|---|---|
| Primary | ITT | Time to the first occurrence of the specific clinical composite outcomes | Investigational Product discontinuation or interruption or dose adjustment | Treatment Policy | Difference in time to the first occurrence of the clinical outcome events between OCA and placebo at end of study |
| | Same as above | Same as above | Study discontinuation due to various reasons or death, or withdrawal or lost to follow up | While on treatment policy | Difference in time to first occurrence of the events including death, and clinical outcomes before study discontinuation, withdrawal or lost to lost to follow. Otherwise, it will be censored. |

IE=intercurrent events; ITT=intent-to-treat; OCA=obeticholic acid

### 2.3.2.2. Secondary Efficacy Endpoints

#### 2.3.2.2.1. Histological Analysis

The analysis of the histological efficacy endpoints at End of Study will be based on the following endpoints analyzed for all post-baseline biopsies.

- Percentage of subjects with Improvement of fibrosis by at least 1 stage with no worsening of NASH
- Percentage of subjects with Resolution of NASH with no worsening of fibrosis

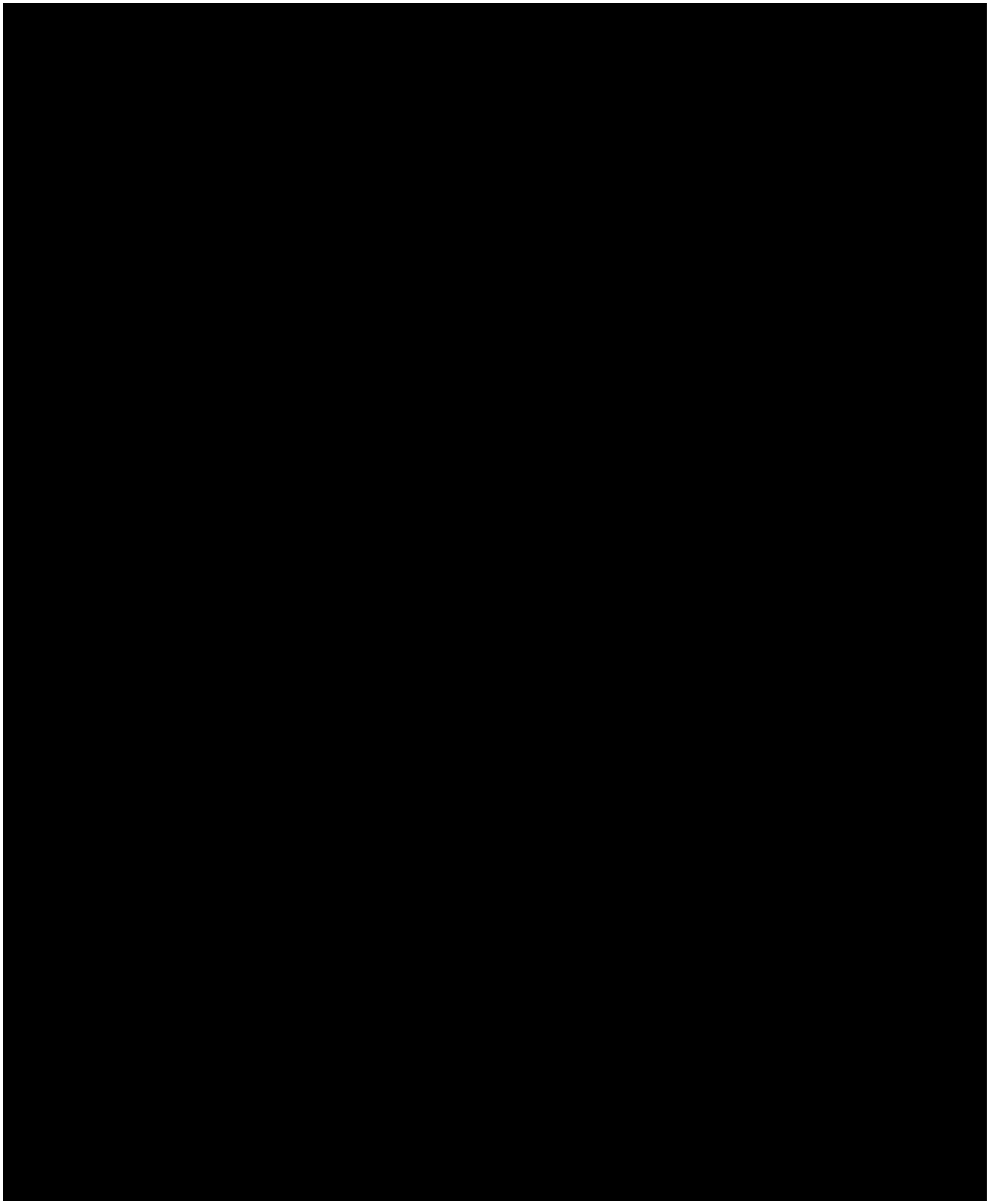
- Percentage of subjects with improvement of fibrosis by at least 1 stage AND/OR resolution of NASH, without worsening of either
- Percentage of subjects with no worsening of fibrosis AND no worsening of NASH
- Percentage of subjects with histological progression to cirrhosis
- Percentage of subjects with improvement of fibrosis by at least 2 stages
- Percentage of subjects with improvement of each key histological feature of NASH by at least 1 point (steatosis, lobular inflammation, and hepatocellular ballooning)
- Percentage of subjects with improvement of NAS by at least 2 points with no worsening of fibrosis
- Percentage of subjects with improvement of fibrosis and resolution of NASH as a composite endpoint and as defined by both endpoints being met in the same subject
- Percentage of subjects with resolution of fibrosis

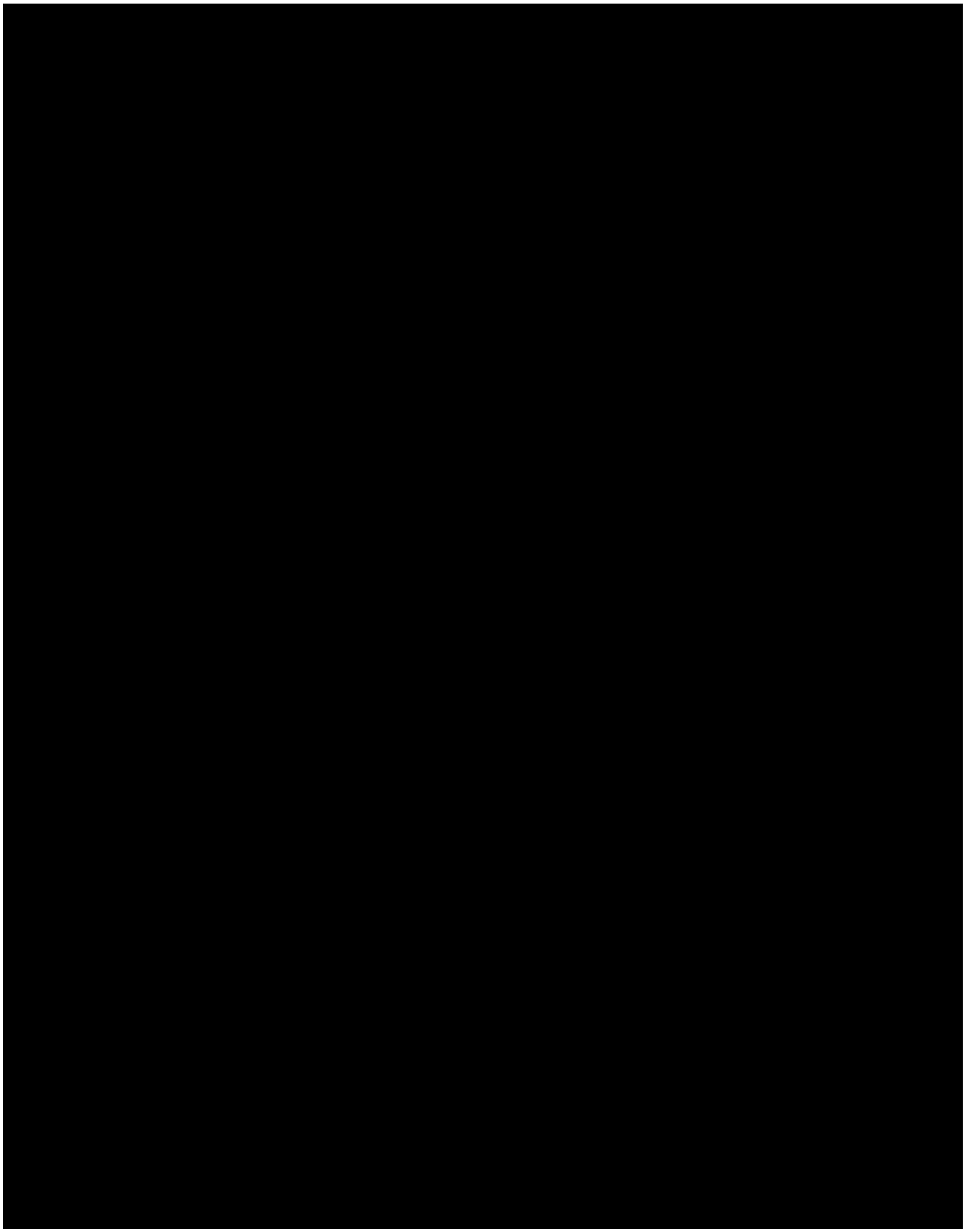
## 2.4. Safety Evaluations at Month 18 and End of Study

Safety-related evaluations include:

- Vital signs
- 12-lead electrocardiograms (including central reading ECGs)
- Clinical laboratory evaluations, including
  - Hematology
  - Serum chemistry
  - Urinalysis
- Lipoprotein concentrations and [redacted]
- TEAEs including AESIs, including pruritus, hepatic disorders, renal disorders (including acute kidney injury), pancreatitis, gallbladder diseases and related complications, dyslipidemia, hyperglycaemia/new onset diabetes, and urolithiases, and cardiovascular events
- Adjudicated cardiovascular events
- Adjudicated hepatic safety events
- Adjudicated acute kidney injury events

[redacted]





## 3.4. Analysis Populations

Analysis populations for the Month 18 Interim Analysis and the End of Study Analysis are summarized in Table 3.

### 3.4.1. Month 18 Interim Analysis Populations

The Month 18 Interim Analysis data cutoff date (DCO) has been pre-specified to include all subjects' data for visits (scheduled or unscheduled) occurring on or before that date. The Interim Analysis Cohort (original Month 18 IA Cohort) will include all randomized subjects who received at least one dose of investigational product by the DCO. The original Month 18 IA Cohort will include all fibrosis stages (stages 1, 2 and 3).

The following analysis populations will be evaluated and used for presentation and analysis of Month 18 Interim Analysis data:

- The Safety Population will include the entire IA Cohort, all fibrosis stages (stages 1, 2 and 3), who received at least 1 dose of investigational product (OCA or placebo) by the DCO. Treatment assignment will be based on the treatment actually received.
- The Full Efficacy Analysis Population will include all original Month 18 IA Cohort subjects, randomized by 15 July 2017, including all fibrosis stages (stages 1, 2, and 3), who received at least 1 dose of investigational product. The selected randomization cut-off date ensures an adequate sample size for the efficacy assessment, considering that subjects randomized by 15 July 2017 would have reached their Month 18 biopsy visit by December 2018 based on 72 week treatment period. Treatment assignment will be based on the randomized treatment.
- The Intent-to-Treat (ITT) Population will include a subset of the Full Efficacy Analysis Population, limited to those subjects with fibrosis stage 2 or stage 3. Treatment assignment will be based on the randomized treatment. The ITT Population will be the primary efficacy analysis population.
- The modified ITT (mITT) Population will include all subjects from the ITT Population, with the exception of a subgroup of subjects who discontinued treatment (before the Month 18 Visit and without an end of treatment biopsy) between 13 Sep 2017 and 19 Dec 2017 for the following reasons: withdrawal of consent, lost to follow-up, and due to other reasons. The observed increase in the number of discontinuations during the above-mentioned period was an unintended consequence of 1) safety communications eliciting concern from subjects and physicians and 2) increased subject and site burden due to the safety amendment. Treatment assignment will be based on the randomized treatment. The mITT population will serve as a supportive analysis population.
- The Per Protocol (PP) Population will include all subjects from the ITT Population who completed at least 15 months of treatment, had a Month 18/ end of treatment biopsy, were on investigational product for at least 30 days immediately preceding the biopsy, and did not have any major protocol deviation (Section 5.5) that could

potentially affect the efficacy conclusions. Treatment assignment will be based on the randomized treatment.

- The PK Population will include all subjects who received OCA and have at least one confirmed fasted analyzable sample. Subjects must have been fasting for approximately 8 hours before the visit and must not have any major protocol deviations that potentially affect exposure levels. The PK Population will be used for OCA PK analysis.

### 3.4.2. End of Study Analysis Populations

- The Safety Population will include all randomized subjects who received at least 1 dose of investigational product (OCA or placebo). Treatment assignment will be based on the treatment actually received.
- The Full Efficacy Analysis Population will include all randomized subjects, all fibrosis stages (stages 1, 2 and 3), who received at least 1 dose of investigational product. Treatment assignment will be based on the randomized treatment.
- The Intent-to-Treat (ITT) Population will be a subset of the Full Efficacy Analysis Population, limited to those subjects with fibrosis stage 2 or stage 3. Treatment assignment will be based on the randomized treatment. The ITT Population will be the primary efficacy analysis population. The Per Protocol (PP) Population will include all subjects from the ITT Population who did not have any major protocol deviation (Section 5.5) that could potentially affect the efficacy conclusions and at least 80% compliance with investigational product (Section 10.1). Treatment assignment will be based on the randomized treatment.
- The PK Population will include all subjects who received OCA and have at least one confirmed fasted analyzable sample. Subjects must have been fasting for approximately 8 hours before the visit and must not have any major protocol deviations that potentially affect exposure levels. The PK Population will be used for OCA PK.

**Table 3: Analysis Populations for the Month 18 Interim Analysis and End of Study Analysis**

| Population | Fibrosis Stage | Description | |
|---|---|---|---|
| | | Month 18 Interim Analysis | End of Study |
| Safety | 1,2 and 3 | Randomized subjects who received at least one dose of investigational product by the DCO (IA Cohort) | Randomized subjects who received at least one dose of investigational product. |
| Full Efficacy Analysis | 1,2 and 3 | All subjects, randomized by 15 July 2017, all fibrosis stages (stages 1, 2 and 3), who received at least 1 dose of investigational product. | Randomized subjects who received at least one dose of investigational product. |
| ITT | 2 and 3 | Subset of Full Efficacy Analysis Population | Subset of Full Efficacy Analysis Population |
| mITT | 2 and 3 | Subset of ITT Population which excludes treatment discontinuation (before the Month 18 Visit and without an end of treatment biopsy) between 13 Sep 2017 to 19 Dec 2017 due to withdrawal of consent, lost to follow up and other reasons | Not applicable |
| Per Protocol | 2 and 3 | Subset of ITT Population who completed at least 15 months of treatment, had a Month 18/ end of treatment biopsy, were on investigational product for at least 30 days just preceding the biopsy, and did not have any major protocol deviation | Subset of ITT Population with no major protocol deviation and at least 80% compliance with investigational product (Section 10.1) |

# 4. GENERAL STATISTICAL METHODS

Individual subject data obtained from electronic case report forms (eCRFs), local laboratories, central laboratories, external sources, and any derived data will be presented in data listings by subject. All data listings that contain an evaluation date will contain a relative study day. Pre-treatment and on-treatment study days are numbered relative to the day of the first dose of investigational product which is designated as Day 1. The preceding day is Day -1, the day before that is Day -2, etc.

For categorical variables, summary tabulations of the number and percentage of subjects within each category (with a category for missing data) of the parameter will be presented. Percentage calculations will be based on non-missing data unless otherwise specified.

In response analysis [yes/no] of histological endpoints, the comparison between treatment groups will be performed using CMH test stratified by the randomization strata (type 2 diabetes at enrollment [yes/no] and use of TZDs/glitazones or vitamin E at baseline [yes/no]).

For continuous variables, the number of subjects, mean, standard deviation (SD), standard error of the mean (SEM), median, 25$^{th}$ and 75$^{th}$ quartiles, minimum, and maximum values will be presented.

Longitudinal summaries of continuous variables will present absolute values, change, and percentage change from baseline at each post-baseline visit.

For continuous endpoints satisfying normality assumptions, change from baseline and percentage change from baseline over time will be analyzed using mixed-effect repeated measures (MMRM) model with treatment, baseline, visit, visit by treatment interaction and stratification factors to be included in the model. An unstructured covariance model will be used. If the computational algorithm fails to converge, an appropriate covariance structure will be selected. Estimates of least-square (LS) means, standard errors will be presented by treatment group. Unless indicated otherwise 95% CIs will be provided for efficacy endpoints by treatment group.

Time-to-event data will be summarized using Kaplan-Meier (KM) methodology using 25$^{th}$, 50$^{th}$ (median), and 75$^{th}$ percentiles with associated 2-sided 95% CIs, as well as the percentage of censored observations. KM estimates will be plotted as a "survival curve" for each treatment group, with the number at risk identified. In addition, time-to-event data will be summarized with descriptive statistics for those subjects with an event. Subjects who do not experience an event will be censored. Subjects who do not experience an event will be censored at the date of last contact for clinical outcome endpoints and will be censored at the earliest date of last dose date+30 days, end of study, death date or date cut-off date for treatment-emergent adverse events (TEAEs). The censor date for time to resolution events will be the earliest date of end of study, death date or date cut-off date.

## 4.1. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software Version 9.4 or higher unless otherwise noted. Medical history and AEs will be coded using Medical Dictionary for Regulatory Activities (MedDRA). Concomitant medications will be coded using World Health Organization (WHO) Drug Dictionary.

## 4.2. Baseline Definitions

The baseline value for statistical analyses of quantitative parameters (except lipoprotein evaluations and markers of glucose metabolism) is defined as the mean of all available evaluations, at the individual subject level, prior to the first administration of investigational product, unless otherwise specified. If there is only one evaluation prior to the first administration of investigational product, then the available data from this evaluation will be used as the baseline value. Quantitative parameters include, but are not limited to, lab parameters (except lipoprotein evaluations and markers of glucose metabolism), electrocardiogram measures, exploratory biomarkers, pharmacodynamic parameters, and vital signs.

The baseline value for analyses of lipid parameters and markers of glucose metabolism is defined as the last fasted evaluation prior to the first administration of investigational product, unless otherwise specified.

The baseline value for analyses of qualitative parameters (eg, normal/abnormal), cardiovascular risk assessments, and patient-reported outcomes is defined as the last evaluation prior to the first administration of investigational product. Baseline definition of each assessment is provided in Appendix A.

## 4.3. Visit Windows

Visit windows will be defined around the target date of the visit as planned in the protocol. Day 1 is the first day of receiving investigational product (OCA or placebo). Month 18 (Week 72) biopsy window is Week 60 to Week 96. For all assessments, other than biopsy, each visit window will have the target date as the center and the lower bound is the midway between the target date of this window and the preceding one. The upper bound is the midway between the target date of the current window and the following one. If more than one assessment is collected within a window, then the assessment closest to the target date will be selected. If more than one assessment with the same distance from the target date, then the last assessment will be used. "Last" will be based on sorting by date, time, record number/sequence number. Appendix B shows the visit windows of each visit up to Month 78.

## 4.4. Examination of Subgroups

There will be 3 types of analyses:

1. Subgroup analyses for the primary and key secondary efficacy endpoint using the ITT, mITT, and PP Populations will include the following factors:
    a. Baseline diabetes status: Yes, No
    b. BMI ( $< 30$ kg/m$^2$, $\geq 30$ kg/m$^2$ and $< 35$ kg/m$^2$, $\geq 35$ kg/m$^2$ )
    c. Baseline Vitamin E or TZD use: Yes, No
    d. Age categories: <65, ≥ 65 years and <75, ≥75 years
    e. Gender: Male, Female
    f. Race: White, Non-White (Black or African American, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native)
    g. Geographic Region: North America, Europe, and the Rest of the World (ROW); including Asia as a subgroup
    h. Baseline Fibrosis Score: F1, F2 and F3
    i. Baseline NAS: (<6, ≥6) (Only for the NASH primary efficacy endpoint)
    j. Baseline and concomitant statin medication use (no concomitant use, prior and concomitant use, new concomitant use)
    k. Baseline and concomitant use of anti-diabetic medications (no concomitant use, prior and concomitant use, new concomitant use)

2. Subgroup analyses for safety (selected safety endpoints) will be generated using the Safety Population and will include the following factors:
   a. Baseline diabetes status; Yes, No
   b. Baseline Fibrosis Score: F1, F2 and F3
   c. Age categories: <65, ≥65 years, ≥75 years
   d. Gender: Male, Female
   e. Baseline in BMI (< 30 kg/$m^2$, ≥ 30 kg/$m^2$ and < 35 kg/$m^2$, ≥ 35 kg/$m^2$ )
   f. Baseline cardiovascular medical history: Yes, No
   g. Baseline Framingham (FRS) risk score: ≤10, >10 - ≤20, >20
   h. Baseline eGFR
   i. Baseline and concomitant statin medication use (no concomitant use, prior and concomitant use, new concomitant use)
   j. Baseline and concomitant medication use of antidiabetic medication (no concomitant use, prior and concomitant use, new concomitant use)
   k. Baseline and concomitant medication use of antihypertensive medication (no concomitant use, prior and concomitant use, new concomitant use)
   l. Quartiles of the mean HDL values after first dose of investigational product for adjudicated cardiovascular event
   m. Quartiles of the mean LDL values after first dose of investigational product for adjudicated cardiovascular event
   n. Baseline gallbladder (ultrasound): Cholecystisis (yes vs no)
3. The change from baseline in the following parameters will be analyzed by response to the primary endpoints and treatment group. These summaries will be based on subjects with Month 18 biopsy.
   a. Liver biochemistry: ALT, AST, ALP, and GGT
   b. Lipids: LDL, HDL, VLDL and total cholesterol

# 5. STUDY POPULATION

## 5.1. Subject Disposition

Subject disposition will be tabulated by treatment group and overall and will include the number randomized, the number randomized but not treated, the number treated in total, the number of subjects in the specified analysis population, the number of subjects who completed Month 18 Visit, the number of subjects who completed Month 18 Visit with a biopsy, the number of subjects who withdrew from investigational product but remained in the study, the number of subjects who withdrew from study prior to completing the study and reason(s) for withdrawal, and the number of subjects who were contacted for follow-up.

Disposition summaries will be presented for each analysis population and will be tabulated by treatment group and overall.

Subject enrollment by study site will be tabulated by treatment group and overall.

A by-subject data listing of study completion information including the reason for premature study withdrawal, if applicable, will be presented.

## 5.2. Demographics and Baseline Characteristics

Demographic variables will include the following:

- Age and groups (<65 years, ≥65 years, ≥75 years) at the time of informed consent
- Gender
- Race
- Ethnicity
- Geographic Region: North America, Europe, and ROW (including Asia as a subgroup)

Other baseline characteristics will include the following:

- Liver fibrosis stage (NASH CRN scoring criteria) by central read and consensus read as available.
- Liver fibrosis stage (Modified Ishak scoring criteria)
- Hepatocellular Ballooning by central read and consensus read as available.
- Lobular Inflammation by central read and consensus read as available.
- Steatosis by central read and consensus as available
- Nonalcoholic fatty liver disease activity score (NAS: <6, ≥6)
- Weight (kg)
- Height (cm)
- BMI ($kg/m^2$): Continuous and Categorical: ( <30 $kg/m^2$, ≥30 $kg/m^2$ , <35 $kg/m^2$, ≥35 $kg/m^2$ Diabetes Status (Yes/No)

- Diabetes Status (Yes/No)
- Hemoglobin A1c (HbA1c; ≤5.6%, 5.7-6.4%, ≥6.5%)
- Low density lipoprotein (LDL: mg/dL) (≥100 mg/dL, ≥190 mg/dL)
- Triglycerides (mg/dL) (≥150 mg/dL)
- High density lipoprotein (HDL; mg/dL) (<40 mg/dL [men] or <50 mg/dL [women])
- Baseline Concomitant Medication Use
  - Lipid-lowering medications, including Statins
  - Statins only
  - Antihypertensive medications
  - Antidiabetic medications
  - Thiazolidinedione (TZD) only
  - Vitamin E
  - Anticoagulants/antiplatelets
  - Bile acid sequestrants
- Baseline Caffeine Intake
- Baseline Alcohol Consumption as assessed by AUDIT Questionnaire (AUDIT Score ≤7, >7)
- Baseline Smoking Habits and Status
- [redacted]

Liver biochemistry, liver function and estimated glomerular filtration rate (eGFR) results at baseline will be summarized as continuous variables and as frequencies for categorizations. The upper limit of normal (ULN) for liver functions suggested by FDA will be used: ALT – 30 U/L, AST – 35 U/L, Total Bilirubin – 1.25 mg/dL, Conjugated (Direct) Bilirubin – 0.3 mg/dL, ALP – 120 IU/L.

- eGFR (chronic kidney disease epidemiology collaboration [CKD-EPI] calculation) and groups (G1/Normal, G2, G3)
- ALT and AST: ≤ upper limit of normal (ULN), >ULN to ≤3x ULN, >3x ULN
- Total Bilirubin and Conjugated (Direct) Bilirubin: ≤ULN, >ULN
- ALP and GGT: ≤ULN, >ULN
- Model for end stage liver disease (MELD) Score

Demographics and Baseline characteristics will be summarized and presented by treatment group and overall for the, ITT, mITT, Full Efficacy Analysis, PP, Safety, and PK Populations.

For continuous measures, number of subjects, mean, SD, median, Q1, Q3, minimum, and maximum will be presented. For categorical measures, number and percentage of subjects will be presented. No inferential statistical comparisons will be performed.

All demographic and Baseline characteristic data will be presented in by-subject data listings.

## 5.3. Baseline NASH Disease Characteristics

Baseline NASH disease characteristics will be summarized using data collected from the NASH Disease History eCRF. Variables from the NASH Disease History eCRF include the following:

- Age at NASH diagnosis
- Duration of NASH in years at time of informed consent

Baseline NASH disease characteristics will be summarized and presented by treatment group and overall for the ITT, mITT, Full Efficacy Analysis, PP, Safety, and PK Populations. For these continuous measures, the number of subjects, mean, SD, median, Q1, Q3, minimum, and maximum will be presented. No inferential statistical comparisons will be performed.

Baseline NASH disease characteristics will be presented in by-subject data listings.

## 5.4. Medical and Cardiovascular Medical History

Verbatim terms on electronic case report forms (eCRFs) will be mapped to preferred terms (PT) and system organ classes using Medical Dictionary for Regulatory activities (MedDRA).

Medical history will be summarized by SOC, PT, and treatment group using the Safety Population. Summaries will be ordered by descending order of incidence of SOC and PT within each SOC.

The following cardiovascular (CV) variables from the CV Medical History eCRF will be summarized by preferred term:

- Types and number of CV conditions/events
- Types and number of CV treatment procedures received

Lastly, separate summaries will be provided by SOC and PT for the number and percentage of subjects with medical history in the following categories:

- Pruritus
- Hepatic Disorders
- Gallbladder disease and Related Complications
- Pancreatitis
- Glycemic control
- Dyslipidemia
- Renal Disorders
- Urolithiases

Medical and CV medical history will be presented in a by-subject data listing.

## 5.5. Protocol Deviations

The Investigator is not permitted to deviate from the protocol in any significant way without notification to the Sponsor (or designee) as described in the protocol.

All eligibility deviations specified in the protocol will be considered major deviations and subjects meeting any of these deviations will not be included in the PP Population. Major on-study protocol deviations that could potentially affect the efficacy or safety conclusions of the study will be identified through programming and identifying selected categories from the list of protocol deviations collected during the study:

- Investigational product dosing, including subjects who receive treatment different from what they were randomized to.
- Taking investigational product after bariatric surgery (gastric bypass)

All major protocol deviations will be summarized and listed by treatment group.

## 5.6. Impact of COVID-19 Pandemic

The percentage and number of subjects with protocol deviations related COVID-19 will be summarized by randomized treatment and by deviation type in Full Efficacy Analysis Population and Safety Population.

The incidence of COVID-19 related TEAEs, serious TEAEs, severe TEAEs, relationship to OCA and fatal TEAE will be summarized by treatment group, by SOC and PT in Safety Population. A listing of subjects with the COVID-19 related TEAEs will be generated.

# 6. EFFICACY ANALYSES

Primary efficacy will be based on the ITT Population. The sections below provide further details on the methods and populations to be used in analyzing each of the efficacy endpoints. All efficacy endpoints will be presented in by-subject data listings.

## 6.1. Adjustments for Covariates

For histological endpoints; primary, secondary, and exploratory endpoints; the comparison between treatment groups will be performed using CMH test stratified by the randomization strata (type 2 diabetes at enrollment [yes/no] and use of TZDs/glitazones or vitamin E at baseline [yes/no]).

For continuous secondary and exploratory endpoints, change from baseline and percentage change from baseline over time will be analyzed using the MMRM model with treatment, baseline, visit, visit by treatment interaction, and stratification factors to be included in the model.

## 6.2. Handling of Subject Discontinuations and Missing Data

Subjects who discontinue investigational product for any reason other than withdrawal of consent are expected to continue in the study until study termination.

In order to determine the effect of missing data on the analysis, efficacy endpoints will be analyzed using different methods of imputation as described below.

### 6.2.1. Quantitative Endpoints

For exploratory efficacy endpoints utilizing an analysis of covariance (ANCOVA), MMRM model, observed cases will serve as the primary analysis.

### 6.2.2. Responder Endpoints

For the primary endpoint analysis, the centrally read screening biopsy will be used as baseline histology data. The Month 18 or post-baseline biopsy that is centrally read as part of the paired read (ie, paired with a new slide from the screening/baseline biopsy tissue to minimize temporal bias) will be compared with the baseline biopsy (screening) to determine efficacy, ie, responder status for primary endpoints as well as other histology efficacy endpoints. Additional analyses may be conducted using the baseline biopsy slide that was used as part of the paired reading for primary endpoints as well as other histology efficacy endpoints.

The protocol requires a Month 18 biopsy even for subjects who discontinued treatment but are on-study (in addition to the end of treatment biopsy, if applicable, by the protocol). For each post-baseline biopsy, a paired reading will be performed. If a subject has more than one paired reading by the Month 18 Visit, only one paired reading will be reported in Electronic Data Capture (EDC) based on the biopsy preceded by the longest (and least interrupted) investigational product exposure.

The biopsy consensus based read methodology by a three pathologists panel will be implemented for all baseline and all subsequent histology data using whole slide image (WSI) technology. The stained slides were digitized (scanned by ICON and other vendor) and the WSI were sent to PathAI (vender) to be read by pathologists using the consensus method using their research platform. In cases where a previously stained slide was either missing OR could not be read by the Consensus panel, a new slide was requested to be stained, scanned and read. A two-stage approach will be utilized to reach a consensus. At stage one, all three pathologists at each panel will independently read each subject's slide image and enter the results in the database. If the two pathologists matched the score for a specific component, it is chosen as the consensus score for the component. If the scores from all 3 pathologists for a specific parameter are discordant, a single consensus score will be provided by a joint panel read at stage two and flagged as Joint panel decision (refer to histology manual).

Panel A of 3 pathologists will read images of slides stained with H&E to characterize NAS (ballooning, inflammation, steatosis) score and Panel B pathologists will read images of slides stained with Trichrome to characterize NAS CRN fibrosis score. The endpoint regarding the improvement or worsening of fibrosis stage and NASH will be derived based on the results from the consensus values at baseline and any post-baseline liver biopsy collection.

For the analyses of histological efficacy endpoints, in which subjects are classified as either responders or non-responders (binary outcome), any subject who discontinues from the study prior to the Month 18 biopsy Visit and does not have a post-baseline biopsy assessment will be considered a non-responder. Biopsies collected for subjects who discontinued treatment before the Month 18 visit will be included in the Month 18 Interim Analysis of histological endpoints regardless of the timing of the biopsy.

Similar conventions and statistical methodologies will be used for analyses of other exploratory histological-based responder analyses. Responder endpoint analyses will be carried out for the primary and secondary endpoints using the ITT, mITT, Full Efficacy Analysis Population and PP Populations.

A complete summary and description of the analysis plan and populations for histological data based on the consensus method can be found in SAP Addendum V2 (dated 25 Apr 2022).

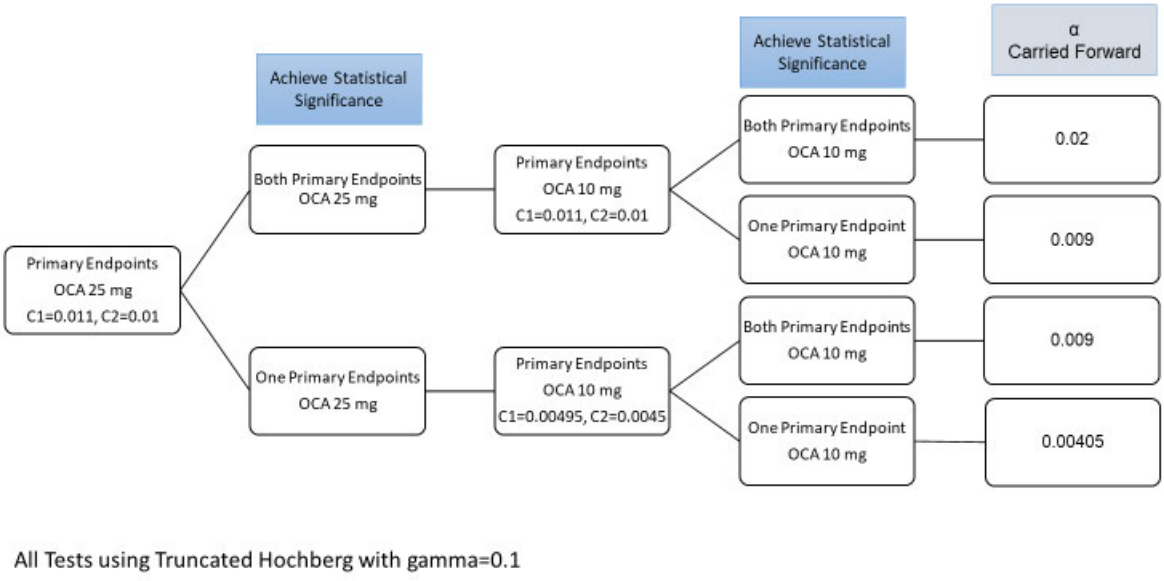
## 6.3. Control of Type I Error

### 6.3.1. Histological Primary Endpoints:

The 2-sided Type I error (alpha) allocated to all testing in this single study will be 0.05. The histological endpoints at the Month 18 Interim analysis will be performed with an alpha level of 0.02. The primary clinical outcomes composite endpoint will be tested with an alpha level of 0.03 which may be augmented by recycled alpha from the Month 18 Interim Analysis.

As shown in Figure 2, the inferential testing will start with the two primary endpoints in comparing the OCA 25 mg dose group versus placebo using the Truncated-Hochberg procedure (Gamma=0.1) and Type I error of 0.02.

**Figure 2: Statistical Testing for the Month 18 Interim Analysis**




## I) Scenario 1: Both Primary Endpoints Achieve Statistical Significance at the OCA 25 mg Dose Level

a. **Primary Endpoints at the OCA 25 mg Dose Level:** Placebo and OCA 25 mg groups will be compared with respect to the fibrosis and NASH primary efficacy endpoints, using the Truncated-Hochberg procedure, with critical values of 0.011 and 0.01 against which the larger and smaller P-value are to be compared respectively. If the larger P-value is less than 0.011 then both primary efficacy endpoints achieve statistical significance in favor of the OCA 25 mg group. Consequently, the full alpha fraction is preserved and will be carried to compare placebo and OCA 10 mg groups with respect to the primary endpoints using the Truncated-Hochberg procedure (Gamma = 0.1).

b. **Primary Endpoints at the OCA 10 mg Dose Level:** Placebo and OCA 10 mg groups will be compared, using the Truncated Hochberg procedure, with respect to the fibrosis and NASH primary efficacy endpoints with critical values of 0.011 and 0.01 against which the larger and smaller P-value are to be compared respectively.

c. **Key Secondary Endpoint at the OCA 25 mg Dose Level:** Placebo and OCA 25 mg groups will be compared with respect to the key secondary endpoint sequentially after the primary endpoint at OCA 10 mg dose level.

d. **Key Secondary Endpoint at the OCA 10 mg Dose Level:** Placebo and OCA 10 mg groups will be compared with respect to the key secondary endpoint sequentially after the key secondary endpoint at OCA 25 mg dose level.

**II) Scenario 2: One Primary Endpoint Achieves Statistical Significance at the OCA 25 mg Dose Level**

a. **Primary Endpoints at the OCA 25 mg Dose Level:** If the larger P-value is greater than 0.011 but the smaller P-value is < 0.01 then only one of the two primary efficacy endpoints achieves statistical significance in favor of the OCA 25 mg group. Preserved alpha of 0.009 will be carried to compare placebo and OCA 10 mg groups with respect to the primary endpoints using the Truncated Hochberg procedure (Gamma = 0.1).

b. **Primary Endpoints at the OCA 10 mg Dose Level:** Placebo and OCA 10 mg groups will be compared, using the Truncated Hochberg procedure, with respect to the fibrosis and NASH primary efficacy endpoints with critical values of 0.00495 and 0.0045 against which the larger and smaller P-value are to be compared respectively.

c. **Key Secondary Endpoint at the OCA 25 mg Dose Level:** Placebo and OCA 25 mg groups will be compared with respect to the key secondary endpoint sequentially after the primary endpoint at OCA 10 mg dose level.

d. **Key Secondary Endpoint at the OCA 10 mg Dose Level:** Placebo and OCA 10 mg groups will be compared with respect to the key secondary endpoint sequentially after the key secondary endpoint at OCA 25 mg dose level.

### 6.3.2. End of Study Primary Endpoints:

For the analyses of the clinical outcomes composite endpoint at the End of Study, OCA doses will be compared to placebo in a sequential closed gate-keeping manner using Type I error of 0.03.

- First placebo and OCA 25 mg will be compared with respect to the clinical outcomes composite endpoint. If the comparison achieves statistical significance at the 2-sided alpha level in favor of OCA 25 mg, then
- Placebo and OCA 10 mg will be compared with respect to the clinical outcomes composite endpoint.

The possible scenarios of testing the End of Study, depending on Month 18 Interim Analysis results, are provided in Table 4.

**Table 4: End of Study Type 1 Error**

| Month 18 Interim Analysis | | | | End of Study | | |
|---|---|---|---|---|---|---|
| OCA 25 mg | | OCA 10 mg | | Testing for Clinical Outcomes Endpoint | | |
| Primary Endpoints | | Primary Endpoints | | | | |
| α | Achieve Statistical Significance | α | Achieve Statistical Significance | α Clinical Outcomes | Both Dose Groups Hierarchal * | Only OCA 10 mg** |
| 0.02 | Both | 0.02 | Both | 0.05 | OCA 25 mg →10 mg at α=0.05 | OCA 10 mg at α=0.05 |
| | | | Only one | 0.039 | OCA 25 mg →10 mg at α=0.039 | OCA 10 mg at α=0.039 |
| | Only one | 0.009 | Both | 0.039 | OCA 25 mg →10 mg at α=0.039 | OCA 10 mg at α=0.039 |
| | | | Only one | 0.03405 | OCA 25 mg →10 mg at α=0.03405 | OCA 10 mg at α=0.03405 |

α for Clinical Outcomes endpoint is 0.03 + carried alpha from the Month 18 Interim Analysis depending on the statistical success of primary endpoints at OCA 25 mg and OCA10 mg dose level
* Following Month 18 Interim Analysis, both OCA arms are continued for Clinical Outcomes.
** Following Month 18 Interim Analysis, only the OCA 10 mg arm is continued for Clinical Outcomes.

## 6.4. Interim Analyses and Data Monitoring

### 6.4.1. Interim Analysis of Clinical Outcomes

The clinical outcome events adjudicated by independent hepatic outcome committee will be closely monitored in a blinded manner until 100% of the required events have been captured. The projections for the total number of events in three groups (placebo, OCA 10mg, and OCA 25mg) will be generated and reviewed following the Performance Standard Plan. When 288 events in all three groups are expected to be observed based on the modeling projections, it is assumed that the targeted 190 events for 66% IA analysis in two groups (placebo and OCA 25) will have accrued. This will trigger the 66% interim analysis evaluations as presented in Section 3.2.

One interim analysis of the clinical composite outcomes endpoint, including all-cause mortality and liver-related clinical outcomes will be conducted when 66% of the 288 clinical outcomes events are accrued in the placebo and OCA 25 mg groups using the Lan-DeMets O'Brien-Fleming boundaries for efficacy (Reboussin 2000). Further details on the boundaries and alpha spent in the interim analysis are provided in Section 3.2.

### 6.4.2. Data and Safety Monitoring Committee

An independent DMC will include hepatologists, physicians, epidemiology/cardiology expert(s), and statistician(s) who will not be involved in the study as Investigators, adjudication committee members, or consultants will have oversight over the study conduct as defined by the DMC charter.

The DMC will meet approximately quarterly, and at least every 6 months, at scheduled meetings and ad hoc meetings, as appropriate, to review safety and efficacy data as well as the adjudication assessments from the adjudication committees as specified in the charter. Based on the unblinded review of these data, the DMC will advise the Sponsor on the validity and scientific merit of continuing the study.

Written minutes of both the open and closed DMC sessions will be prepared. The closed minutes will be made available to the appointed Sponsor representatives only after the database is locked.

Data flow, access, process, and disclosure will be described in a DAP and Unblinding Plan.

### 6.4.3. Adjudication Committees

All suspected liver-related clinical outcomes and major adverse cardiovascular events (MACE) events that occur after administration of the first dose of investigational product will be reviewed by the following blinded adjudication committees, depending on event type, before inclusion in any analysis. The committees and event types they are responsible for adjudicating are as follows:

- Cardiovascular Adjudication Committee: Adjudicates all suspected MACE, including all deaths following the CV Adjudication Charter.
- Hepatic Outcomes Adjudication Committee: The role of the HOC is to review and adjudicate all deaths and suspected liver-related clinical outcomes following the pre-specified criteria described in the HOC Charter.
- Renal Adjudication Committee (RAC) – The role of the RAC is to review potential Acute Kidney Injury (AKI) events and adjudicate an event as AKI following the pre-specified criteria included in the RAC Charter.
- Hepatic safety Adjudication Committee (HSAC) - The role of the HSAC is to review potential hepatic injury triggers to determine whether they meet the pre-specified criteria for hepatic injury as described in the HSAC Charter.

Each adjudication committee will have a charter that documents the reporting of events by the study site, definition of the suspected events to be adjudicated, supply of source documentation to the committee, entire data flow and process from committee membership, review of the events by the committee, reporting of the final assessment, and the working procedures of the committee.

All potential liver-related clinical outcome events, hepatic safety, renal and major cardiovascular adverse events (MACE) will be reviewed by adjudication committees before inclusion in any analysis. The adjudication committee members will be independent hepatologists, nephrologists (cardiologists for MACE) not involved in the study as investigators, DMC members, or consultants. All adjudication committee members will remain blinded to treatment group throughout the study. The data coordinating center will remove all subject identification on the eCRF data and source documentation, as specified in the charter.

### 6.4.4. Primary Efficacy Analysis

#### 6.4.4.1. Primary Efficacy Analysis at Month 18

The primary efficacy analysis at Month 18 will compare OCA to placebo. The primary efficacy analysis will be conducted using the ITT Population.

The primary endpoints for the Month 18 Interim Analysis are:

- Improvement of fibrosis by at least 1 stage with no worsening of NASH (no worsening of hepatocellular ballooning, no worsening of lobular inflammation and no worsening of steatosis)
- Resolution of NASH with no worsening of fibrosis

The primary efficacy analysis at OCA 25 mg dose will test the following hypotheses using the ITT Population:

- $H_{01}$: The percentage of subjects with fibrosis improvement by 1 stage or more, no worsening of hepatocellular ballooning, no worsening of lobular inflammation, and no worsening of steatosis is equal between placebo and OCA 25 mg groups.
- $H_{11}$: The percentage of subjects with fibrosis improvement by 1 stage or more, no worsening of hepatocellular ballooning, no worsening of lobular inflammation, and no worsening of steatosis is different between placebo and OCA 25 mg groups.
- $H_{02}$: The percentage of subjects with NASH resolution and no worsening of fibrosis is equal between placebo and OCA 25 mg groups.
- $H_{12}$: The percentage of subjects with NASH resolution and no worsening of fibrosis is different between placebo and OCA 25 mg groups.

The primary efficacy analysis at OCA 10 mg dose will test the following hypotheses using the ITT Population:

- $H_{01}$: The percentage of subjects with fibrosis improvement by 1 stage or more, no worsening of hepatocellular ballooning, no worsening of lobular inflammation, and no worsening of steatosis is equal between placebo and OCA 10 mg groups.
- $H_{11}$: The percentage of subjects with fibrosis improvement by 1 stage or more, no worsening of hepatocellular ballooning, no worsening of lobular inflammation, and no worsening of steatosis is different between placebo and OCA 10 mg groups.
- $H_{02}$: The percentage of subjects with NASH resolution and no worsening of fibrosis is equal between placebo and OCA 10 mg groups.
- $H_{12}$: The percentage of subjects with NASH resolution and no worsening of fibrosis is different between placebo and OCA 10 mg groups.

The primary efficacy analyses at Month 18 will compare placebo and each OCA dose, adjusting for multiplicity as described in Section 6.3 and using a CMH test stratified by the randomization strata (diabetes at enrollment [yes/no] and use of TZDs or vitamin E at baseline [yes/no]).

These analyses, in addition to the sensitivity analyses described in Section 6.2.3, will also be conducted using the mITT, PP, and Full Efficacy Analysis Populations.

#### 6.4.4.2. Primary Efficacy Analysis at End of Study

The primary efficacy endpoint at the End of Study will be the time to first occurrence of one of the following hepatic adjudicated events (and progression to cirrhosis based on post baseline biopsy including Month 48 biopsy) post-randomization:

- Death (all-cause)
- MELD score ≥15
- Liver transplant
- Hospitalization (as defined by a stay of 24 hours or greater) for onset of:
  - Variceal bleed
  - Hepatic encephalopathy (as defined by a West Haven score of ≥2)
  - Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis)
- Ascites secondary to cirrhosis and requiring medical intervention (eg, diuretics or paracentesis)
- Histological progression to cirrhosis

The hypotheses testing for the primary efficacy endpoint at the End of Study is:

- $H_{01}$: The time to first occurrence of any of the hepatic adjudicated events for all-cause mortality and liver-related clinical outcomes are equal between placebo and OCA 25 mg groups.
- $H_{11}$: The time to first occurrence of any of the hepatic adjudicated events for all-cause mortality and liver-related clinical outcomes are different between placebo and OCA 25 mg groups.
- $H_{02}$: The time to first occurrence of any of the hepatic adjudicated events for all-cause mortality and liver-related clinical outcomes are equal between placebo and OCA 10 mg groups.
- $H_{12}$: The time to first occurrence of any of the hepatic adjudicated events for all-cause mortality and liver-related clinical outcomes are different between placebo and OCA 10 mg groups.

The primary efficacy analysis will be conducted using the ITT Population, only including fibrosis stage 2 and 3 subjects. Placebo and each OCA dose will be compared separately using a log rank test stratified by the randomization stratification factor. Kaplan-Meier (KM) estimates of the distribution of the time-to-event will be tabulated and graphed by treatment group. The tabulation will include the KM estimate of the 25$^{th}$, 50$^{th}$ (median), and 75$^{th}$ percentiles and corresponding 95% confidence intervals (CIs), if the medians can be estimated. The number and percentage of subjects censored and with events will be presented. The hazard ratio and 95% CI

will be estimated using a Cox regression model stratified by randomization strata to estimate the magnitude of the effect.

### 6.4.5. Secondary Efficacy Analyses at Month 18 and End of Study

The secondary efficacy analysis will be conducted using the ITT Population. The secondary endpoints for the Month 18 Interim Analysis are:

Key Secondary Endpoint:

- Percentage of subjects with improvement of fibrosis by at least 1 stage AND/OR resolution of NASH, without worsening of either

Secondary Endpoints:

- Percentage of subjects with no worsening of fibrosis AND no worsening of NASH
- Percentage of subjects with progression to cirrhosis
- Percentage of subjects with improvement of fibrosis by at least 2 stages
- Percentage of subjects with improvement of each key histological feature of NASH by at least 1 point (steatosis, lobular inflammation, and hepatocellular ballooning)
- Percentage of subjects with improvement of NAS by at least 2 points with no worsening of fibrosis
- Percentage of subjects with improvement of fibrosis and resolution of NASH as a composite endpoint and as defined by both endpoints being met in the same subject Percentage of subjects with resolution of fibrosis.

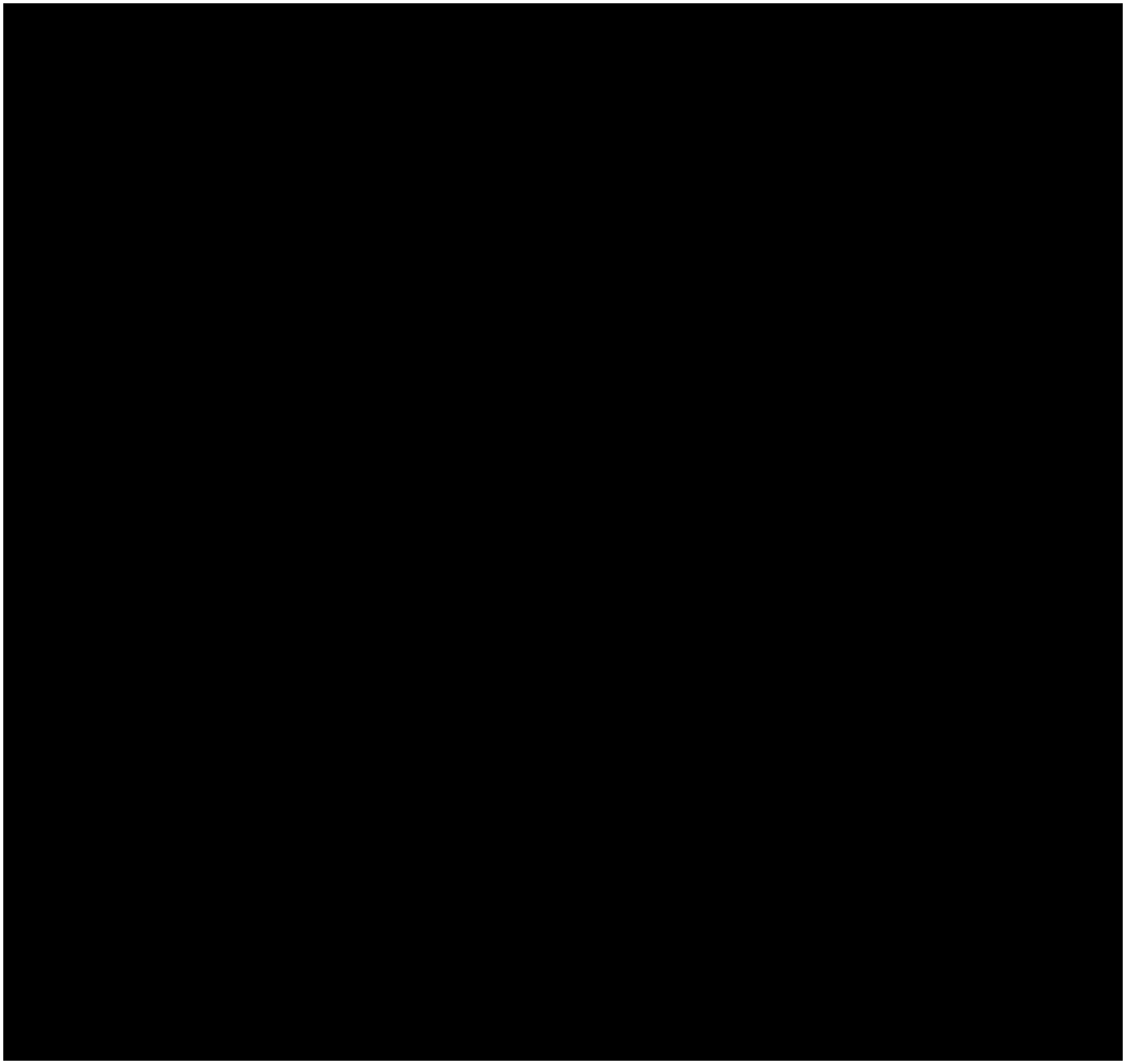
The analyses will compare placebo and each OCA dose separately (10 mg and 25 mg), using a CMH test stratified by the randomization strata (diabetes at enrollment [yes/no] and use of TZDs or vitamin E at baseline [yes/no]).

These analyses will also be conducted using the mITT, PP and Full Efficacy Analysis Populations.

At the End of Study Analysis, all Month 18 histological endpoints will be analyzed.

[redacted]

[redacted]



### 6.4.6.10. Change from Baseline in MELD

MELD score is a scoring system for assessing the severity of chronic liver disease. Higher scores on MELD indicate worse disease severity.

MELD during treatment will be compared to baseline values to determine if there was an increase in MELD score, and this information will be summarized by treatment group and visit. MELD score, change and percentage change from baseline values will be summarized and analyzed as described in the General Statistical Methods (Section 4).

Individual factors contributing to the MELD score will be presented in a listing.

### 6.4.7. End of Study Time to Event Analyses

In addition to the primary analysis (Section 6.4.4.2) and secondary analyses (Section 6.4.5) and exploratory analyses (Section 6.4.6), the following analyses will be performed at the End of Study.

#### 6.4.7.1. Time to Each Component of the Clinical Outcome Endpoint

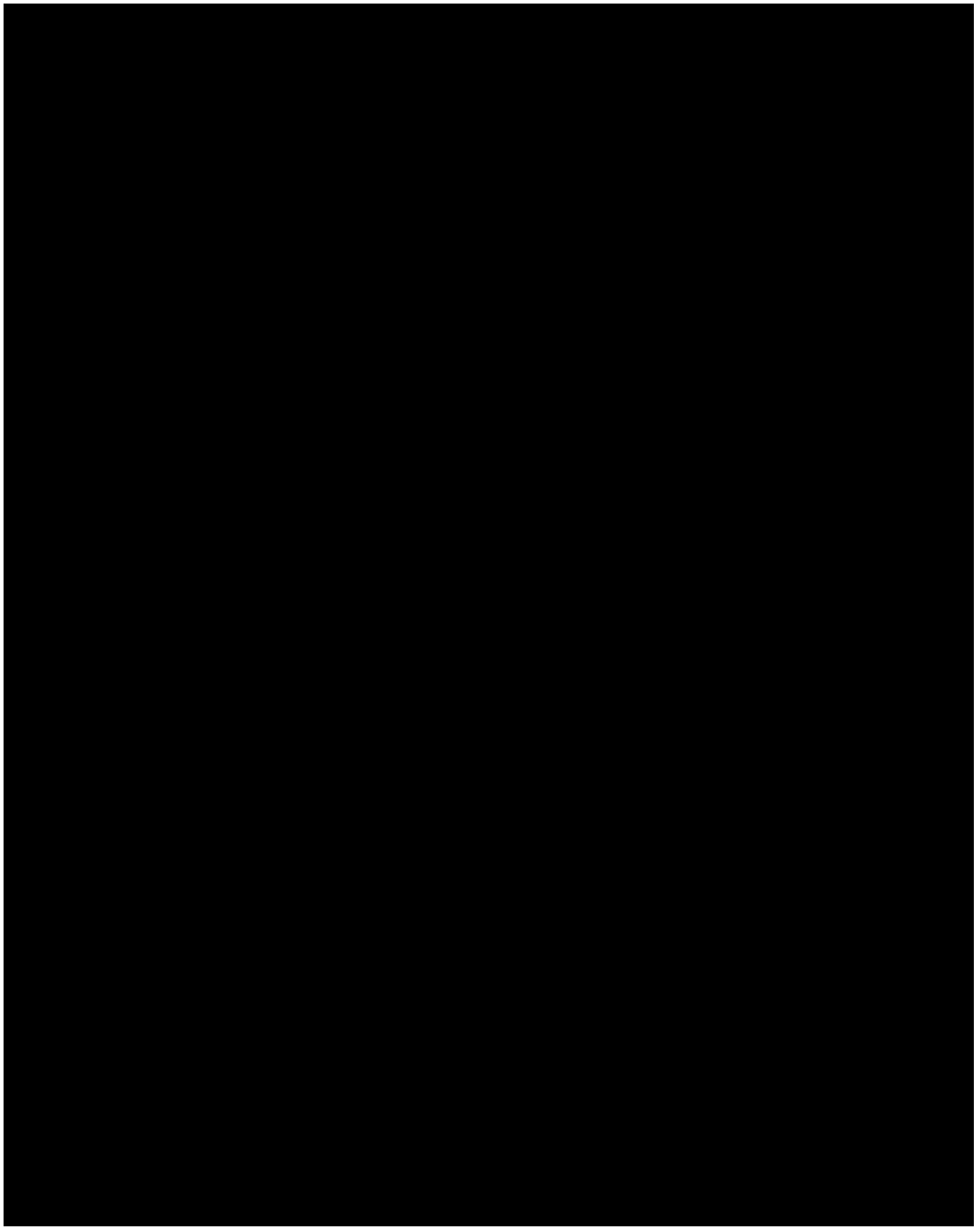
Time to each of the following components of the clinical outcome endpoint and time to HCC will be analyzed by treatment group using ITT and Full Efficacy Analysis Populations:

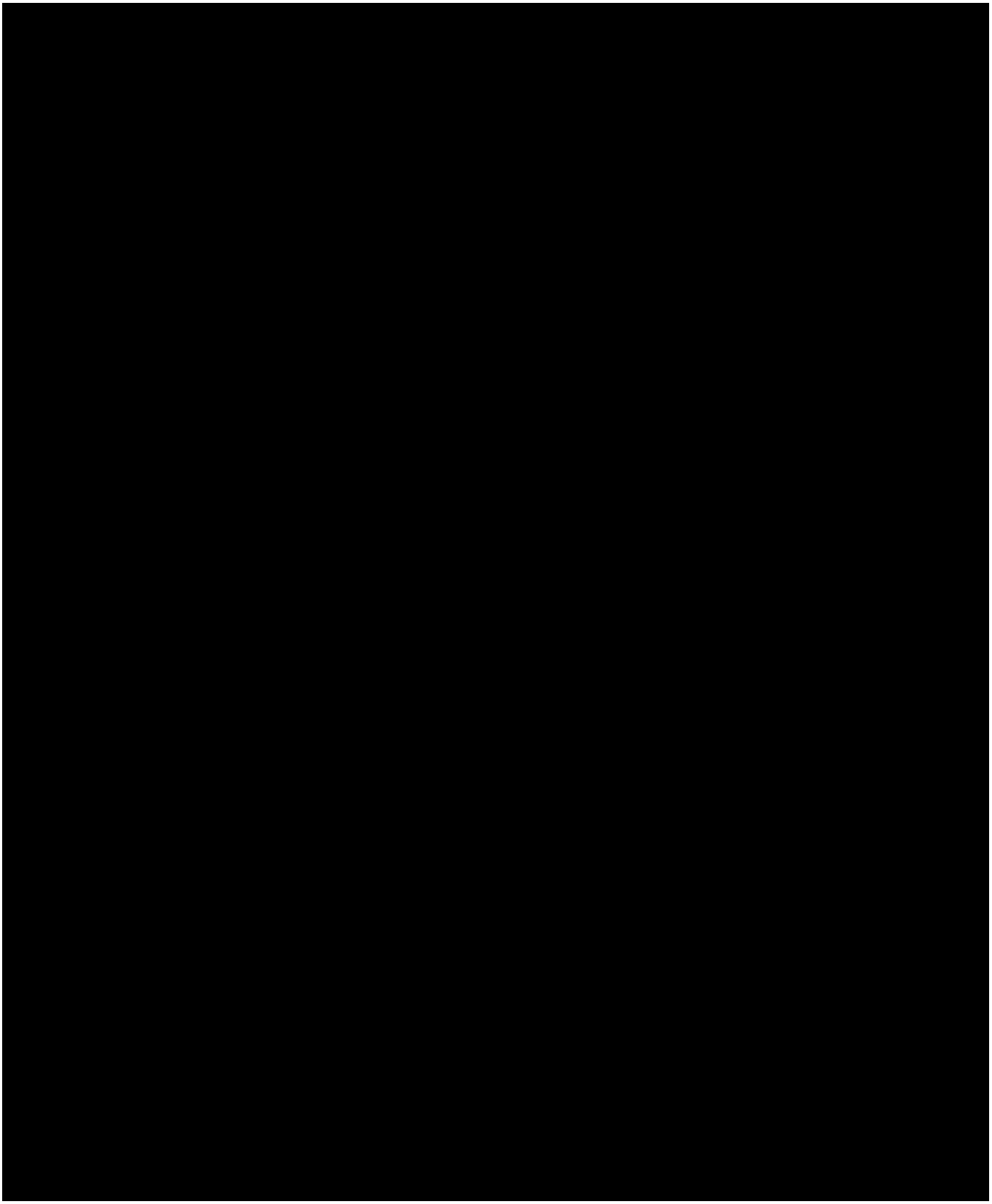
- Death: date of death
- Liver transplant: date of first liver transplant
- MELD score ≥15: date of first occurrence of MELD ≥15
- Hospitalization (as defined by an inpatient stay of 24 hours or greater) for new onset or recurrence of: variceal bleed, encephalopathy (as defined by a West Haven grade ≥2), spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis): date of earliest hospitalization admission.
- Ascites secondary to cirrhosis and requiring medical intervention: date of earliest ascites start date as an adverse event that required medical intervention.
- Histological progression to cirrhosis: date of biopsy

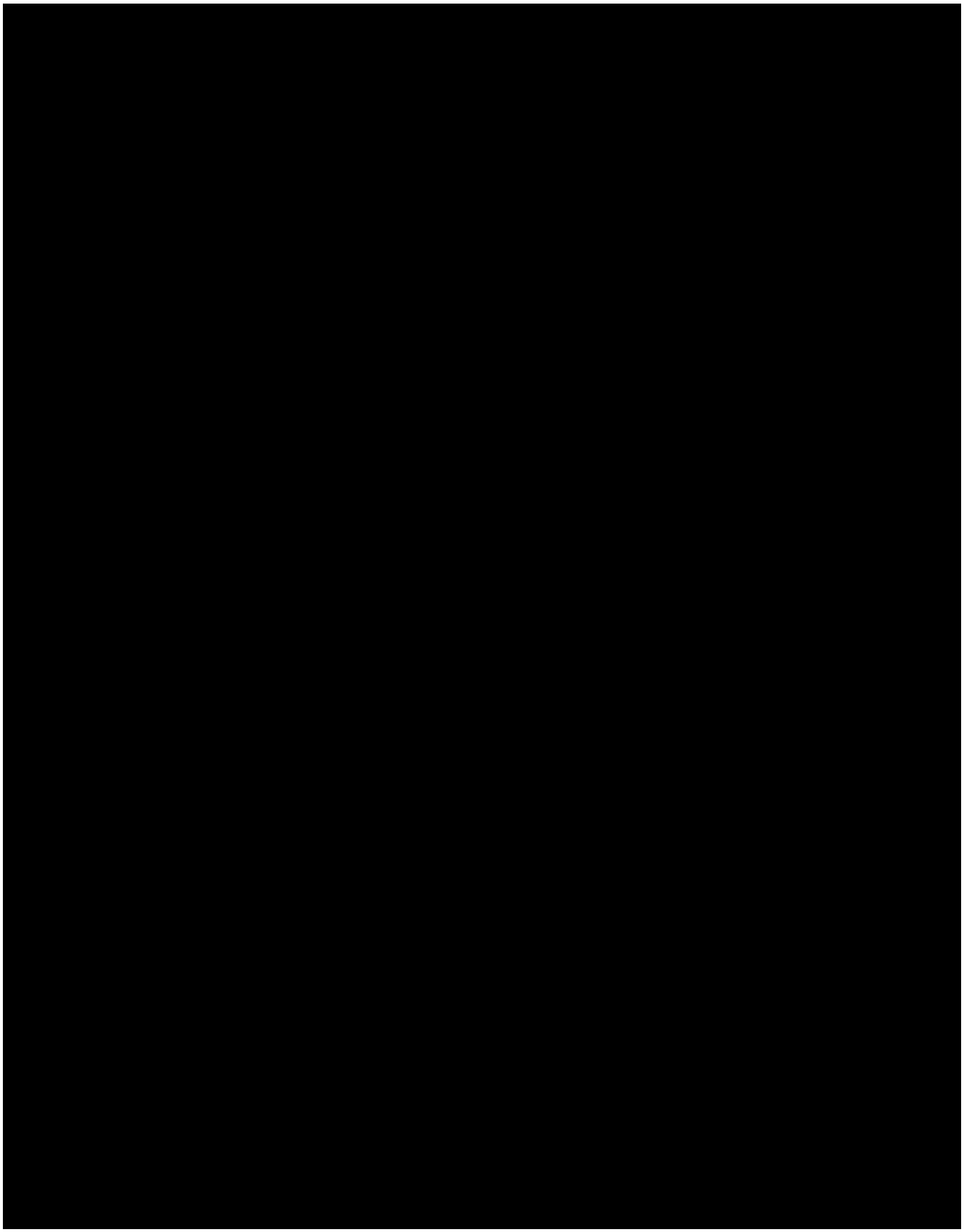
Only adjudicated events (and biopsy confirmed progression to cirrhosis) will be included in the analyses. The time to event of each of the components of the clinical outcome endpoint are defines as follows:

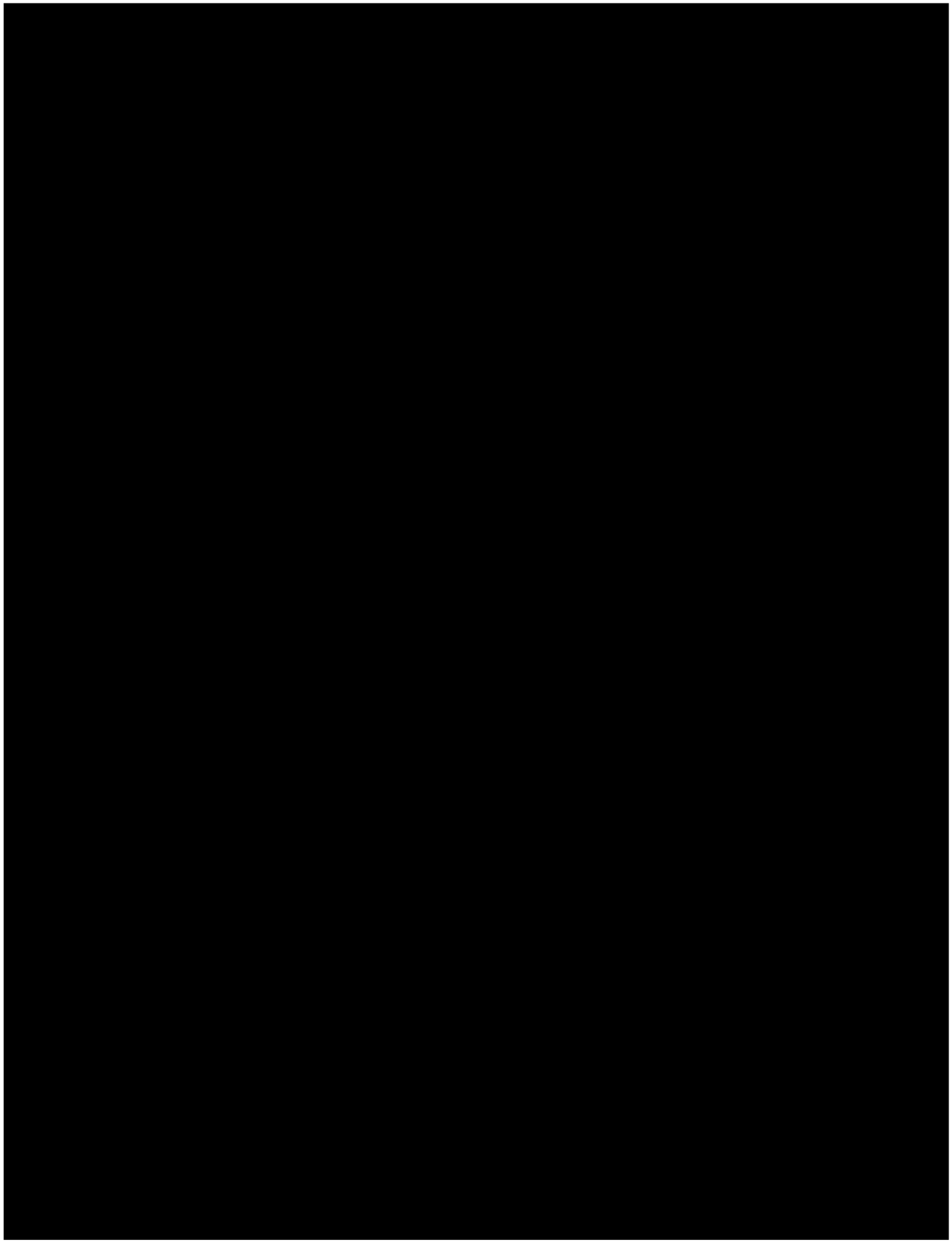
- Time to death (all-cause), first MELD score ≥15 and first liver transplant will be calculated in days as the date of the event minus the date of randomization plus 1. Subjects who do not experience an event will be censored at the date of last contact.
- Time to the first occurrence of onset of ascites (secondary to cirrhosis and requiring medical intervention), will be calculated in days as the first time for onset of the ascites minus the date of randomization plus 1.
- Time from randomization to the date of biopsy confirming histological progression to cirrhosis will be calculated in days as the date of the diagnosis minus the date of randomization plus 1. Subjects who do not experience an event will be censored at the date of last contact.
- Time to the first occurrence of hospitalization for new onset or recurrence of variceal bleed, encephalopathy, spontaneous bacterial peritonitis will be calculated in days as the first hospitalization admission date/time for new onset or recurrence of the event minus the date of randomization plus 1. The duration between the hospitalization admission date/time and discharge date/time (discharge date/time – admission date/time) must be greater than or equal to 24 hours (1 day). Subjects who do not experience an event will be censored at the date of last contact.

KM estimates of the time-to-event will be tabulated and graphed by treatment group. The tabulation will include the KM estimate of the medians and corresponding 95% CIs, if the medians can be estimated. The number and percent of subjects censored and with events will be presented. All summaries of incidence will include the associated exact binomial 95% confidence Interval (CI). Listing of adjudicated events, including progression to cirrhosis will be provided. At the Month 18 Interim Analysis, a listing of adjudicated events will be prepared by DMC vendor without unblinding the Sponsor as described in the DAP.









[redacted]

# 10. SAFETY ANALYSES

Safety analyses will be conducted using the Safety Population including by treatment groups and total OCA (both OCA 25 mg and OCA 10 mg combined). If a subject is randomized to placebo and subsequently receives OCA, each TEAE (Section 10.3) will be attributed to the administered OCA dose.

No inferential comparison of safety endpoints will be performed, unless otherwise specified.

There will be 2 sets of safety summaries presenting the safety profile of the following populations:

1. The Safety Population as defined in Section 3.4
2. The ITT population as defined in Section 3.4

Frequencies of adverse events will be split into 2 periods: pre-Covid 19 (occurring on or before 30 Jan 2020) and post-Covid 19 (occurring after 30 Jan 2020). The data for both periods will be presented side-by-side by treatment group for the safety population.

Selected safety summaries will be prepared including all subjects in the mITT Efficacy Population. Summaries and listings will be presented by actual treatment received.

## 10.1. Extent of Exposure

The extent of exposure will be summarized both continuously using descriptive statistics and categorically (n and % for ≤4 weeks, >4 and ≤12 weeks, >12 and ≤24 weeks, >24 and ≤36 weeks, >36 and ≤48 weeks, >48 and ≤72 weeks, >72 and ≤96 weeks, >96 weeks ≤120 weeks; >120 weeks ≤144 weeks, >144 weeks ≤168 weeks, >168 weeks ≤192 weeks). After Month 48 (Week 192) until the end of the study, exposure will be presented by 48 weeks periods.

Extent of exposure (days) to investigational product will be calculated as follows:

- Exposure to investigational product = {[(Date of last investigational product dose – Date of 1st investigational product dose)+1] – Total duration of temporary investigational product discontinuation, interruption, drug holidays and/or missed doses}

The duration of each incidence of temporary investigational product discontinuation will be calculated as follows:

- Duration of temporary discontinuation, interruption, and/or missed doses of investigational product = (Date of restart of investigational product – Date of temporary discontinuation, interruption, drug holidays and/or missed doses of investigational product) + 1.

The total duration of temporary investigational product discontinuation is the sum duration of temporary discontinuation, interruption, and/or missed doses of investigational product over each incidence of discontinuation / interruption.

Total investigational product (mg) exposed to subject will be calculated by adding the doses taken by a subject during the study and will be summarized using descriptive statistics.

A summary of subjects who had an at least one temporary investigational product interruption, discontinuation, and/or missed doses (including drug holiday) during the study will be provided in terms of frequency count (n) and percentages (%) and mean duration of the interruptions/missed doses. This summary will be presented by reason of interruption as specified in the eCRF.

Subject's overall compliance (%) with investigational product will be calculated as follows:

- (# of tablets consumed during study / # of tablets expected to be consumed during study) * 100

where # of tablets consumed during study = # of tablets dispensed- # of tablets returned or

- total investigational product (mg) exposed to subject/ total investigational product (mg) planned

where total investigational product planned = (End date of IP exposure – Start date of IP exposure + 1)

Subject compliance with investigational product will be summarized by treatment group using descriptive statistics. For discontinued subjects, the expected tablets to be consumed will be calculated based on the time of discontinuation.

All exposure data will be presented in a by-subject data listing.

## 10.2. Prior and New Concomitant Medications

Verbatim terms on eCRFs will be mapped to Anatomical Therapeutic Chemical (ATC) class and Preferred Name using the WHO Drug Dictionary Enhanced.

Pretreatment concomitant medications are those medications with start and stop prior to the initial dose of investigational product also classified as no concomitant medication. Prior concomitant medications are those medications started prior to and continued after the initial dose of investigational product. New concomitant medications are those medications that were started after the initial dose of investigational product. If it cannot be determined whether the medication was a new concomitant medication due to a partial start or stop date or if the medication is taken on the same date as the initial dose of investigational product, then it will be counted as a new concomitant medication.

No concomitant, prior, and new concomitant medications will be summarized separately for each treatment group by WHO ATC level 2, WHO ATC level 4, and preferred name using the Safety Population. These summaries will present the number and percentage of subjects using each medication. Subjects may have more than one medication per ATC level and preferred name. At each level of subject summarization, a subject is counted once if he/she reported one or more medications at that level. Each summary will be ordered by descending order of incidence of ATC level and preferred name within each ATC level.

Extent of exposure of newly initiated use (after the 1st dose) of vitamin E, thiazolidinediones, bile acid sequestrants (BABR), statins and other lipid-lowering medications, will be calculated as follows:

> Exposure to the concomitant medication = {[(Date of last dose –Date of 1st dose)+1] – Total duration of temporary interruptions}

Missing start and end dates will be imputed as described in Appendix A. Extent of exposure will be summarized by treatment group as described in the General Statistical Methods (Section 4).

## 10.3. Adverse Events

AE data will be collected from the time that signed informed consent is obtained until the subject fully completes her/his study participation of the study. AEs will be reported and categorized using the Common Terminology Criteria for Adverse Events (CTCAE, Version 4.03).

All AE summaries will be restricted to TEAEs, which are defined as any AEs that newly appear, increase in frequency, or worsen in severity following randomization up to 30 days from the last dose date of permanent investigational product discontinuation up to new data cut-off date. If it cannot be determined whether the AE is treatment emergent due to a partial onset date, then it will be counted as such. Verbatim terms on eCRFs will be mapped to preferred terms (PT) and system organ classes (SOC) using MedDRA version 18.1 or later version.

Each AE summary will be displayed by treatment group. Summaries that are displayed by SOC and PT will be ordered by descending order of incidence (OCA 25 mg, followed by OCA 10 mg, followed by placebo) of system organ class and preferred term within each system organ class. Summaries of the following types will be presented:

- Overall summary of TEAEs
- Subject incidence of TEAEs and the total number of entries by MedDRA SOC and PT.
- Subject incidence of TEAEs and the total number of entries by PT in descending order.
- Subject incidence of serious TEAEs and the total number of entries by MedDRA SOC and PT.
- Subject incidence of TEAEs leading to investigational product withdrawal by MedDRA SOC and PT.
- Subject incidence of TEAEs leading to study discontinuation by MedDRA SOC and PT.
- Subject incidence of TEAEs by MedDRA SOC, PT, and highest severity (CTCAE grade). At each level of subject summarization, a subject is classified according to the highest severity if the subject reported 1 or more events. AEs with missing severity (CTCAE grade) will be considered grade 3 (severe) for this summary.
- Subject incidence and total number of entries of CTCAE grade 3 or higher TEAEs by MedDRA SOC and PT. At each level of subject summarization, a subject is

classified according to the highest severity if the subject reported 1 or more events. AEs with missing severity (CTCAE grade) will be considered grade 3 (severe) for this summary.

- Subject incidence and total number of entries of related TEAEs by MedDRA SOC and PT. At each level of subject summarization, a subject is classified according to the closest relationship to investigational product if the subject reported 1 or more events. TEAEs with a missing relationship will be considered related for this summary.
- Subject incidence and total number of entries of related CTCAE grade 3 or higher TEAEs by MedDRA SOC and PT. At each level of subject summarization, a subject is classified according to the highest severity if the subject reported 1 or more events. AEs with missing severity (CTCAE grade) will be considered grade 3 (severe) for this summary. TEAEs with a missing relationship will be considered related for this summary.
- Subject incidence of TEAEs occurring by MedDRA SOC and PT with frequency at least 5% within any treatment group.
- Subject incidence of TEAEs by MedDRA SOC, PT, and closest relationship to biopsy (Related/Not Related). At each level of subject summarization, a subject is classified according to the closest relationship to biopsy if the subject reported 1 or more events. AEs with a missing relationship will be considered related for this summary.
- Subject incidence of TEAEs leading to death and the total number of entries by MedDRA SOC and PT

The following listings will be presented by treatment group and subject:

- All TEAEs
- Serious TEAEs (This is a subset of the TEAEs where serious is marked as “Yes”)
- CTCAE Grade 3 or higher TEAEs (This is a subset of TEAEs where severity is marked as CTCAE grade 3, 4, or 5)
- Study Drug Related TEAEs (This is a subset of the TEAEs where relationship to study drug marked as “Definite,” “Probable,” or “Possible”)
- Biopsy-related TEAEs (This is a subset of the TEAEs where relationship to biopsy marked as “Definite”, “Probable”, or “Possible”)
- TEAEs leading to Study Drug Withdrawal or Study Discontinuation (This is a subset of the TEAEs where Action Taken with Study Treatment is checked as “Drug Withdrawn” or “Subject Discontinued from Study)
- TEAEs leading to death (This is a subset of the TEAEs where outcome is indicated as “Fatal” or the CTCAE grade is 5)

## 10.4. Adverse Events of Special Interest

The following treatment-emergent AESIs will be summarized:

- Pruritus
- Hepatic
- Cardiovascular
- Dyslipidemia
- Gallbladder disease and related complications
- Pancreatitis
- Renal
- Urolithiases
- Hyperglycaemia/New onset diabetes mellitus

Each subcategory of interest will be defined using Standardized MedDRA Queries (SMQs) and/or custom MedDRA queries, as specified in the subsequent sections.

AESI analyses will include all of the following:

- Incidence of treatment-emergent AESIs by PT
- Incidence of treatment-emergent AESIs by PT and severity
- Incidence of serious treatment-emergent AESIs by PT
- Incidence of treatment-emergent AESIs leading to IP withdrawal by PT
- Incidence of related treatment-emergent AESIs by PT
- Incidence of treatment-emergent AESIs leading to death by PT
- Exposure-adjusted incidence of AESIs by PT
- Exposure-adjusted incidence of serious AESIs by PT
- Exposure-adjusted incidence of AESIs by baseline diabetes status
- Exposure-adjusted incidence of severe (Grade 3+) AESIs by PT
- KM statistics and plot for time to first onset of treatment-emergent AESI
- KM statistics and plot for time to onset of first severe (Grade 3+) or serious treatment-emergent AESI
- KM statistics and plot for duration of first treatment-emergent AESI
- KM statistics and plot for duration of most severe treatment-emergent AESI
- Percentage of subject-days with treatment-emergent AESIs per subject exposure year (SEY), by severity

All summaries noted above will include crude incidence rates for each level of summation. For overall incidence summary of AESI, analyses will be performed separately for AESI collected on-treatment only and during study.

**Exposure-adjusted incidence rate** (EAIR) difference (with corresponding 95% Cis) will be calculated for the number of subjects with any event by subcategory of interest as well as for each PT. EAIR will be calculated to correct for differences in investigational product exposure by using person-time in the denominator. Adjusted incidence per 100 SEY is defined as the number of subjects with an event for whom person-time is available divided by the total SEY for each treatment group and multiplied by 100. If a subject experiences the event of interest, the subject's SEY will be calculated as (first event onset date, minus the first dose date plus 1) divided by 365.25 days/year. If a subject does not experience the event of interest, the subject's SEY will be calculated as (earliest date of the last dose date plus 30 days or study discontinuation date or death date or date cut-off date minus the first dose date plus 1) divided by 365.25 days/year. One SEY is the equivalent of one subject exposed to investigational product for one year. Two subjects who are exposed to investigational product for half a year together contribute one SEY. The total SEY of a treatment group is the sum of the SEY for each subject in that treatment group. Exposure-adjusted incidence rates will be calculated for each treatment group. Exposure-adjusted incidence rate difference and the corresponding 95% CI will be derived from Poisson regression model with treatment as explanatory variables.

The time to onset of first treatment-emergent AESI, including onset of first treatment-emergent AESI with different categories, will be estimated using the KM method by calculating the time from the first dose date to the first occurrence of a specific category of treatment-emergent AESI. Subjects who do not experience an event will be censored as described in General Statistical Methods (Section 4).

The duration of first treatment-emergent AESI and most severe treatment-emergent AESI will be estimated using the KM method by calculating the time from the start date of first occurrence of a given treatment-emergent AESI or most severe AESI to its resolution (AE outcome being "recovered/resolved" in the eCRF). If the AESI remains ongoing, subjects will be censored as described in General Statistical Methods (Section 4). Only subjects with the events will be included in the analysis.

KM estimates will be plotted as a "survival curve" for each treatment group, with the number at risk identified. In addition, time-to-event data will be summarized with descriptive statistics for those with an event.

A by-subject data listing will be provided for AEs of special interest (AESI) by category including the Standardized MedDRA query (SMQ) where applicable.

Specific analyses for each AESI will be described in the subsequent sections.

### 10.4.1. Pruritus

Treatment-emergent pruritus, defined as any preferred term within the pruritus Not Elsewhere Classified (NEC) high level term (HLT) or including "prur," will be summarized separately by the MedDRA SOC, treatment group, and PT as a subset of all TEAEs. Additional summaries of

overall pruritus TEAEs will be provided. These summaries will be generated for the Safety and ITT Populations.

For subjects reporting an AE of pruritus, the various strategies for management of pruritus (listed below) will be summarized (per FDA request). Frequencies per treatment group will be reported. In addition, interruption of investigational product and decreased dose frequency will be summarized and presented in the following ways:

1. each subject represented only once,
2. average interruptions per subject,
3. number of interruptions lasting: 1-5 days, 6-10 days, >10 days.

An analysis of treatment-emergent pruritus event days per patient years on study will be performed. The days at each severity grade, including the total days of event, the total patient years on study, and the event days per patient year on study at each severity grade, including the total, will be summarized. Each individual patient year on study will be derived as (the earliest date of last dose date plus 30 days, study discontinuation date, death date and data cut-off date minus the first dose date plus 1) divided by 365.25 days/year.

Strategies for management of pruritus (≤ Grade 2 Severity) include:

- Interruption of investigational product
- Decreased dose frequency of investigational product
- Any medication initiated for pruritus
  - bile acid sequestrants
  - antihistamines
  - corticosteroids
  - other (e.g., naltrexone/naloxone, rifampin/rifampicin)
- None of these strategies (no management).
- Two listings of all treatment-emergent pruritus events will be provided: the standard AE listing, and an additional listing which will include the concomitant medication(s) received, adjustments to investigational product dosing frequency or amount, mandatory investigational product interruption, and/or holiday. The time between each event and its preceding one will be presented per subject.

### 10.4.2. Hepatic Safety Analysis

Hepatic disorders are different and separate from primary efficacy endpoint of hepatic clinical outcomes and are defined as events included in the Hepatic Disorders SMQ, excluding the following sub-SMQs: alcohol related; congenital, familial, neonatal and genetic disorders of the liver; liver infections; and pregnancy-related hepatic disorders.

Summaries will include overall summary of hepatic disorder TEAEs, summary of the incidence of hepatic disorder TEAEs, related TEAEs, TEAEs with CTCAE Grade 3 or higher, TEAEs

leading to treatment discontinuation, TEAEs leading to study discontinuation, SAEs, TEAEs leading to death.

In order to explore the relationship between hepatic disorders and OCA, the following time-to event analyses will be performed:

- EAIR of custom hepatic queries for PTs "Acute hepatic failure" (hepatic failure) and "Ammonia increased" (hepatic injury)
- Incidence of hepatic TEAEs by PT, and baseline fibrosis status
- Incidence of hepatic SAEs by PT, and baseline fibrosis status

Potential drug induced liver injury (DILI) Triggers: Potential cases of DILI in Study 747-303 are identified via a set of biochemical, adverse event (SMQ based) and study visit questionnaire triggers. These cases then undergo clinical review by the Hepatic Safety Adjudication Committee (HSAC) of all available data for each identified subject to determine liver injury status (including causality and severity).

Summary tables for identified cases will present frequencies of triggers by treatment group as well as highest causality and severity association.

Additional DILI analyses will include the following:

- EAIR of DILI cases adjudicated for ALL cases of hepatic injury by severity and treatment arms
- EAIR of DILI cases adjudicated as ≥ possible causality and severity injury by treatment arms
- Incidence of potential DILI by baseline weight (<85kg, 85-100 kg, and >100 kg)
- Incidence of potential DILI by baseline BMI (<30 kg/m2, ≥30 kg/m2 to <35 kg/m2, ≥35 kg/m2)
- Incidence of potential DILI by use of known hepatotoxic medications prior to the event
- Incidence of potential DILI by baseline fibrosis using histology based on original eligibility baseline fibrosis stage.
- Time to onset of first occurrence of adjudicated DILI (Possible, Prob, Highly Likely causality AND ≥ Moderate severity)
- Time to onset of first occurrence of adjudicated DILI (Possible/Prob/Highly Likely causality AND Mild severity)
- Time to onset of first occurrence of adjudicated DILI (Unlikely causality AND >/= Moderate severity)
- Time to onset of first occurrence of adjudicated DILI (Unlikely causality AND Mild severity)

### 10.4.3. Dyslipidemia

Dyslipidemia adverse events are defined as events included in the dyslipidemia SMQ. Analyses to be performed specifically for dyslipidemia events include:

- Incidence of dyslipidemia TEAEs by PT and presence of dyslipidemia medical history (yes, no)
- Incidence of dyslipidemia SAEs by PT and presence of dyslipidemia medical history (yes, no)
- Incidence of dyslipidemia TEAEs by PT and statin use at baseline (no concomitant use, prior and concomitant use, new concomitant use)
- Incidence of dyslipidemia SAEs by PT and statin use at baseline (no concomitant use, prior and concomitant use, new concomitant use)
- Incidence of LDL increase ≥15%
- Incidence of sustained LDL increase ≥15% (defined as present for 2 or more consecutive visits)
- Incidence of dyslipidemia TEAEs preceding any Gallbladder disease and Related Complications TEAE
- KM plot for time to onset of first dyslipidemia TEAE
- Incidence of LDL >100, >130, >190 post baseline (BL) by treatment group.
- Incidence of LDL increase ≥15%

### 10.4.4. Gallbladder Disease and Related Complications

TEAEs of Gallbladder Disease and Related Complications, including cholelithiasis and cholecystitis, are defined as events included in the Gallbladder Related Disorders narrow SMQ and Gallstone Related Disorders narrow SMQ, with additional PTs of Biliary abscess, Biliary sepsis, Biliary tract infection, Gallbladder abscess, Gallbladder empyema, Bile duct necrosis, Bile duct obstruction, Bile duct stenosis, Biliary colic, Cholangitis, Cholangitis acute, Cholangitis chronic, Cholecystocholangitis, and Perforation bile duct.

Specific analyses for gallbladder/gallstone associated events will include:

- Incidence of Gallbladder disease and Related Complications TEAEs by PT and baseline BMI (<30 kg/m2, ≥30 to <35 kg/m2, ≥35 kg/m2)
- Incidence of Gallbladder and Related Complications TEAEs by PT and body weight quartiles
- Incidence of Gallbladder disease and Related Complications TEAEs by PT and baseline diabetes status (defined as history of diabetes mellitus at baseline or on antidiabetic medications at baseline with indication of diabetes mellitus or baseline HbA1c ≥6.5%)
- Incidence of Gallbladder disease and Related Complications TEAEs by PT and presence of Gallbladder disease and Related Complications medical history.

- Number and percentage of subjects who underwent a cholecystectomy
- Number and percentage of subjects who underwent a cholecystectomy on or subsequent to adverse event onset day
- KM plot for time to onset of first Gallbladder disease and Related Complications TEAE
- KM plot for time to onset of cholecystectomy.
- KM plot for time to onset of cholecystectomy on or subsequent to gallstone/gallbladder related event onset day

### 10.4.5. Pancreatitis

Treatment-emergent adverse events of pancreatitis, defined by Acute Pancreatitis narrow SMQ, will be summarized as follows:

- Incidence of pancreatitis TEAEs, by PT and presence of Gallbladder disease and Related Complications medical history (yes, no)
- Incidence of pancreatitis TEAEs, by PT and presence of Pancreatitis (acute or chronic) medical history (yes, no)
- Incidence of pancreatitis TEAEs by PT and presence of Gallbladder disease and Related Complications TEAE (occurring before or simultaneous to Pancreatitis TEAE)
- Incidence of pancreatitis TEAEs by glucagon-like peptide-1 (GLP-1) agonist use (never used, baseline use, initiated during study, baseline use or initiated during study); Note: GLP-1 use initiated during study will only be counted if initiated prior to onset of first pancreatitis TEAE.
- Incidence of pancreatitis TEAEs by shift from baseline to worst post-baseline triglycerides >1000 mg/dL. Note: Only the worst post-baseline triglycerides measurement recorded prior to onset of first pancreatitis TEAE is utilized.
- KM plot for time to onset of first pancreatitis TEAE
- KM plot of pancreatitis cases with on treatment gallbladder Disease or related complications (occurring before or simultaneous to Pancreatitis)

### 10.4.6. Renal Injury

Treatment-emergent renal injury is broadly defined as events included in the Acute Renal Failure SMQ, Chronic Kidney Disease SMQ, Proteinuria SMQ, Renovascular disorders SMQ and Tubulointerstitial disease SMQ. Analyses to be performed specifically for renal events include:

- Incidence of renal TEAEs by PT and presence of renal medical history (yes, no)
- Incidence of renal SAEs by PT and presence of renal medical history (yes, no)
- Incidence of renal TEAEs in subjects with >50% increase in serum creatinine from baseline

- Incidence of renal TEAEs by worst change from baseline eGFR (≥25% reduction from baseline – yes, no); Note: Only the worst post-baseline eGFR measurement recorded prior to onset of first renal TEAE is utilized.
- EAIR specifically for PTs in the broad SMQ "acute renal failure"
- Listing and narratives will be provided for any subjects who had:
  - ≥2x increase in serum creatinine compared to baseline
  - ≥50% decrease in eGFR compared to baseline
  - TEAE of "Nephrolithiasis" (search based on that specific MedDRA PT)

In addition of assessing kidney injury during the trial, the following analyses will be performed, per FDA feedback, to identify a **potential kidney injury**, using the PT terms as noted in Appendix C.

- To assess severity of acute kidney injury (AKI) the following will be provided:
  - Proportion of subjects with TEAE by PT (as noted in Appendix C, Table C1 [List 1]) including risk differences and 95% CIs (using Cochran-Mantel-Haenszel [CMH] method)
  - Proportion of subjects with TEAE by PT (as noted in Appendix C, Table C2 [List 2]) including risk differences and 95% CIs (using CMH method)
- Proportion of subjects who had an action taken due to AKI adverse events including risk differences and 95% CIs (using CMH method)
- Incidence of eGFR decrease ≥25% and ≥50 % from baseline by chronic kidney disease (CKD) stage at baseline based on the standard eGFR categories as defined in Table 7 (Section 10.5.6).
- A listing of subjects who discontinue treatment because of renal dysfunction and their associated last renal lab values while on treatment will be provided by treatment and eGFR decrease.

### 10.4.7. Urolithiases

Treatment-emergent urolithiases is defined as events included in the high-level group term Urolithiases (SOC Renal and Urinary Disorders). Analyses to be performed specifically for urolithiasis events include:

- incidence of urolithiases TEAEs in subjects with >50% increase in serum creatinine from baseline prior to urolithiases.

### 10.4.8. Hyperglycaemia/New Onset Diabetes Mellitus

Hyperglycemia is defined as adverse events included in the hyperglycaemia/new onset diabetes mellitus narrow SMQ. Analyses to be performed specifically for hyperglycemia/diabetes events include:

- Incidence of Hyperglycaemia/new onset diabetes mellitus TEAEs and SAEs by PT and baseline diabetes status (history of diabetes mellitus at baseline or on anti-diabetic medications at baseline with indication for diabetes mellitus or baseline HbA1c ≥6.5%).
- Incidence of Hyperglycaemia/new onset diabetes mellitus SAEs by PT and baseline diabetes status (history of diabetes mellitus at baseline or on anti-diabetic medications at baseline with indication for diabetes mellitus or baseline HbA1c ≥6.5%).
- KM plot for time to onset of first Hyperglycaemia/new onset diabetes mellitus TEAE by baseline diabetes status
- Incidence of initiation of antidiabetic medication in any non-diabetes mellitus subject (no history of diabetes at baseline or not already on medication or HbA1c <6.5% at baseline).
- Incidence of antidiabetic agent intensification (defined as addition of another oral agent/ injectable antidiabetic agent / or insulin) post-baseline in subjects with history of diabetes mellitus or medication for diabetes mellitus at baseline.
- Incidence of treatment emergent HbA1c ≥6.5% by treatment group and time to event post-baseline of HbA1c ≥6.5% by treatment group will be presented.

### 10.4.9. Cardiovascular Risk

#### 10.4.9.1. Adjudicated Cardiovascular Events

Adjudicated cardiovascular events, including Major Adverse Cardiovascular Events (MACE) (defined as death, myocardial infarction, and stroke) and other related events including hospitalization for unstable angina, revascularization procedures, hospitalization for congestive heart failure, transient ischemic attack, and additional events related to adverse cardiovascular outcomes, will be included in the CEC Charter for adjudication. The adjudication of these events will be handled separately from the adjudication of events for assessment of the primary clinical composite outcome endpoint in this study.

Summaries of adjudicated cardiovascular events including MACE will include the overall summary of cardiovascular TEAEs, incidence of TEAEs, incidence of serious TEAEs, annualized adjudicated CV event rate, and a time-to-event analysis as described below. All summaries of incidence will include the associated exact binomial 95% confidence interval (CI). Summaries will be carried out for the Safety Population and ITT Population.

The time-to-event endpoints include:

- Time from randomization to the first confirmed occurrence of cardiovascular death, nonfatal myocardial infarction, nonfatal stroke
- Time from randomization to the confirmed occurrence of cardiovascular death (including fatal myocardial infarction, fatal stroke)
- Time from randomization to the first confirmed occurrence of myocardial infarction (nonfatal or fatal)

- Time from randomization to the first confirmed occurrence of stroke (nonfatal or fatal)
- Time from randomization to the confirmed occurrence of death from any cause
- Time from randomization to the first confirmed occurrence of unstable angina requiring hospitalization (nonfatal or fatal)
- Time from randomization to the first occurrence of coronary revascularization procedure
- Time from randomization to the first occurrence of hospitalization for congestive heart failure
- Time from randomization to the first occurrence of transient ischemic attack
- Time from randomization to the first occurrence of hospitalization for clinically significant arrhythmias
- Time from randomization to the first occurrence of hospitalization for peripheral revascularization procedures

In addition, subjects with any adjudicated cardiovascular event will be presented by the following lipid subgroups:

- Total Cholesterol (TCHOL): Mean and median TCHOL values by visit after first dose of investigational product.
- HDL: Mean and median HDL values by visit values after first dose of investigational product
- LDL: Mean and median LDL values by visit after first dose of investigational product
- LDL/HDL: Mean and median LDL/HDL values by visit after first dose of investigational product
- Apolipoprotein B: Mean and Median apolipoprotein B values by visit after first dose of investigation product

Subgroup analysis by prior CV disease, by CV risk factors (eg, FRS), by baseline type 2 diabetes mellitus (T2DM) status will be performed.

Additional analyses for all adjudicated CV events will include:

- EAIR of CV cases adjudicated for ALL cases of CV by severity and treatment arms
- EAIR of CV cases adjudicated as ≥ possible causality and severity by treatment arms

#### 10.4.9.2. Cardiovascular Adverse Events

CV adverse events are defined as events included in the Embolic and thrombotic events broad SMQ, Ischemic heart disease broad SMQ, or Central nervous system vascular disorders narrow SMQ.

Analyses to be performed specifically for CV events include:

- Incidence of CV TEAEs by SOC and PT
- Incidence of CV TEAEs by PT and presence of CV medical history (yes, no)
- Incidence of CV SAEs by PT and presence of CV medical history (yes, no)
- Incidence of CV SAEs by PT and baseline LDL quartiles
- Incidence of CV SAEs by PT and LDL increase >15% from baseline
- Incidence of CV SAEs by PT and baseline HDL quartiles

[redacted]

#### 10.4.9.4. Serum Chemistry Lipids

Summary statistics will be presented for total cholesterol, LDL, HDL, very low density lipoprotein (VLDL), and triglycerides by visit, as well as worst and last postbaseline value. Plots will also be created for these parameters showing mean values over time. Summary tables will be repeated for the following subgroups:

- Baseline LDL quartiles (Reference categories: <100 mg/dL; 101-103 mg/dL; 131-159 mg/dL; >160 mg/dL)
- Baseline HDL quartiles

Additionally, shift tables from baseline to each post-baseline visit will be presented for HDL and LDL, using the normal and abnormal cutoffs below:

- LDL: normal (<100 mg/dL) and abnormal (≥100 mg/dL)
- HDL: normal (≥40 mg/dL in men, or ≥50 mg/dL in women) and abnormal (<40 mg/dL in men or <50 mg/dL in women)

## 10.5. Clinical Laboratory Evaluations

Laboratory parameters will be summarized in both conventional units and the standard international (SI) system of units. Analysis of efficacy-related liver biochemistry endpoints is specified in Section 6.4.6.3.

The submitted datasets will also include results from local lab evaluations and these values will be flagged in the ADLB dataset with the flag variable LOCALFL="Y". Dataset with conventional units will be labelled using standardized names (e.g. ADLB) and the one with SI units will be labelled by adding "SI" to the file name (e.g. ADLBSI).

All clinical laboratory data will be presented in by-subject data listings.

### 10.5.1. Serum Chemistry, Hematology, and Coagulation

Quantitative serum chemistry (including calculated eGFR), hematology, and coagulation results by treatment group using descriptive statistics at baseline and at each on-study evaluation. The laboratory values, change, and percentage change from baseline, will also be summarized. Baseline is defined as the mean of all available evaluations prior to treatment.

Shifts in CTCAE toxicity grade of laboratory tests from baseline to worst value (scheduled or unscheduled), last value (scheduled or unscheduled), and at each scheduled post-baseline visit will be provided. The toxicity grades are in the CTCAE version 4.03.

Summaries will present the number and percentage of subjects with shifts in laboratory toxicity grade by treatment group.

### 10.5.2. Shifts in CTCAE Toxicity Grade

Shifts in CTCAE toxicity grade of laboratory tests from baseline to worst value (scheduled or unscheduled), last value (scheduled or unscheduled), and at each scheduled post-baseline visit will be provided. The toxicity grades are in the CTCAE Version 4.03.

Summaries will present the number and percentage of subjects with shifts in laboratory toxicity grade by treatment group.

### 10.5.3. Urine Chemistry and Urinalysis

Quantitative urine chemistry and urinalysis results by treatment group using descriptive statistics at baseline and at each on-study evaluation. The lab results and percentage change from baseline will also be summarized. Baseline is defined as the mean of all available evaluations prior to treatment.

### 10.5.4. Liver Laboratory Criteria for Monitoring for Suspected Hepatic Injury or Decompensation

The evaluation of hepatic safety will include the following laboratory parameters: ALT, AST, and ALP, GGT, total bilirubin, direct bilirubin, albumin, INR of prothrombin time, and platelets.

Summary statistics will be presented by visit, including change from baseline and percent change from baseline. In addition, mean values of ALT, AST, ALP, GGT, total bilirubin, direct bilirubin, and albumin will be plotted over time.

Summaries will be repeated by baseline fibrosis stage.

In addition, the hepatic safety parameters noted above will be utilized to present analyses of worst post-baseline elevations. EAIR will be presented for each of the following parameters/criteria at any post-baseline visit:

- ALT (≥ULN, ≥3x ULN, ≥5x ULN, ≥10x ULN, ≥20x ULN, or an increase of 250 U/L from baseline)
- AST (≥ULN, ≥3x ULN, ≥5x ULN, ≥10x ULN, ≥20x ULN, or an increase of 250 U/L from baseline)
- ALP (≥2x ULN, ≥3x ULN, or an increase of 150 U/L from baseline)

- Total Bilirubin (≥2x ULN, ≥5x ULN, ≥8x ULN, or an increase of 2 mg/dL from baseline)
- Direct Bilirubin (≥2x ULN, ≥5x ULN, or an increase of 1 mg/dL from baseline)
- GGT (≥2x ULN)
- INR (≥1.5x ULN, ≥3x ULN, ≥5x ULN

Evaluation of drug-induced serious hepatotoxicity (eDISH) will be conducted to visually compare maximum ALT and total bilirubin values at any post-baseline visit. For subjects with baseline ALT or total bilirubin levels above ULN, eDISH plots will also be generated based on multiples of elevations above the baseline (BL) value (2×BL and 3×BL for ALT, and 2×BL for total bilirubin).

Similar eDISH analysis as described above will be performed for ALP and total bilirubin. Plots based on multiples of elevations above the baseline value will use 2×BL for ALP and 2×BL for total bilirubin.

The summary table as well as visual plots will be based on central lab data. eDISH using maximum post-baseline ALT and total bilirubin will be repeated with central and local lab data.

In addition, EAIR will be presented for each of the following biochemical triggers for potential DILI by treatment group using the worst post-baseline elevations. For subjects with normal baseline liver biochemistry:

- (ALT ≥3x ULN or AST ≥3x ULN) and total bilirubin ≥2x ULN
- ALT ≥5x ULN or AST ≥5x ULN
- ALT ≥8x ULN or AST ≥8x ULN
- (ALT ≥3x ULN or AST ≥3x ULN) and INR >1.5
- ALP ≥2x ULN and direct bilirubin ≥2x ULN
- GGT ≥2x ULN and direct bilirubin ≥2x ULN

For subjects with elevated BL liver biochemistry:

- (ALT ≥3x BL or AST ≥3x BL) and total bilirubin ≥2x BL
- ALT ≥3x BL or AST ≥3x BL
- ALT ≥5x BL or AST ≥5x BL
- (ALT ≥2x BL or AST ≥2x BL) and INR >1.5
- ALP ≥2x BL and ALP >ULN and direct bilirubin ≥2x BL and direct bilirubin >ULN
- GGT ≥2x BL and GGT >ULN and direct bilirubin ≥2x BL and direct bilirubin >ULN

For the above summaries the following triggers will be also considered to allow identification of subjects in whom abnormalities were not detected using elevations relative to baseline:

- Trigger 1: BL ALT + 250 U/L
- Trigger 2: BL AST +250 U/• Trigger 3: BL ALT + 150 U/L and BL total bilirubin + 2 mg/dL
- Trigger 4: BL AST + 150 U/L and BL direct bilirubin + 1 mg/dL
- Trigger 5: BL AST or ALT + 150 U/L and BL total bilirubin + 2 mg/dL
- Trigger 6: BL AST or ALT + 150 U/L and INR >1.5
- Trigger 7: BL AST or ALT + 100 U/L and BL total bilirubin + 2 mg/dL and INR >1.5
- Trigger 8: BL ALP + 150 U/L and BL TB + 2 mg/dL
- Trigger 9: BL ALP + U/L and Baseline direct bilirubin + 1 mg/dL
- Trigger 10: BL total bilirubin + 2 mg/dL
- Trigger 11: BL direct bilirubin + 1 mg/dL
- Trigger 12: INR >1.5 (confirmed with repeat test and not responsive to vitamin K administration)

Any elevation ≥1.5x ULN will be categorized as a "elevated" baseline value. For these analyses, the following ULN thresholds will be used:

- ALT – 30 U/L
- AST – 35 U/L
- Total Bilirubin – 1.25 mg/dL
- Direct bilirubin – 0.3 mg/dL
- ALP – 120 IU/L

A summary of the number (%) of subjects who met the stopping criteria provided in Table 3 of the protocol (see Appendix D) will be presented by treatment group and each of the criteria. A listing of these subjects with the associated laboratory parameters will be presented along with the action taken with the study drug and flags to denote values or subjects that meet the proposed mitigation criteria for DILI.

### 10.5.5. Lipoprotein Evaluations

Serum chemistry-based assessments are the standard safety tool for monitoring subjects' serum lipids. Additional lipoprotein samples, separate from the serum chemistry samples, will be obtained and assayed in addition to the serum chemistry panel lipid assay.

The lipoprotein analyte values, change from baseline, and percentage change from baseline will be summarized by treatment group for the lipoprotein analytes. Baseline is defined as the last fasting assessment before treatment. This analysis will use only samples that have a confirmed fasting of approximately 8 hours or more before their visit.

Analyses of quantitative lipoprotein analytes will be carried out using MMRM as described in the General Statistical Methods (Section 4).

Subgroup analysis by statin use (no concomitant statin use, statin use started prior to first dose of investigational product and statin use started after first dose of investigational product) will be performed. In addition to statin use, similar subgroup analyses will be performed by use of other antihyperlipidemic drugs. Only descriptive statistics will be provided.

### 10.5.6. Renal Function Evaluations

Summary statistics for absolute values, change from baseline, and percent change from baseline to each on-study evaluation visit, as well as worst and last postbaseline values, will be presented for eGFR, serum creatinine, and urinary albumin to creatinine ratio as specified in Table 8. The summary will also be repeated for baseline CKD stage, baseline diabetes status (history of diabetes mellitus at baseline or on antidiabetic medications at baseline with indication for diabetes mellitus or baseline HbA1Cc ≥ 6.5%), and baseline fibrosis stage. The 95% CIs for the difference in mean change will be also presented.

Annualized change in eGFR (mL/min/1.73 m2/yr) will be summarized descriptively. The proportion of subjects that experience rapid eGFR decline (defined as >4 mL/min/1.73 m2/yr) will be presented.

Plots will be created to show mean laboratory (kidney function) data change from baseline over time for eGFR, serum creatinine, and as appropriate other markers of kidney function or injury such as Albumin to creatinine ratio. An additional plot for the percent change from baseline over time will be created only for serum creatinine.

Descriptive statistics for eGFR will be presented for subgroups by baseline CKD stage, baseline diabetes status, and baseline fibrosis stage.

Renal safety parameters serum creatinine and eGFR will also be utilized to present analyses of the worst change from baseline at any post-baseline visit, including risk differences and 95% CIs (using CMH method). The following change from baseline categories will be analyzed:

- Creatinine (≥1.5 x BL, ≥2.0 x BL, ≥3.0 x BL)
- eGFR (≥25% decrease from BL, ≥50% decrease from BL, ≥75% decrease from BL)

A listing will be created to display subjects with any post-baseline serum creatinine value ≥2.0 x BL or eGFR ≥50% decrease from BL.

A summary will also be provided based only on the last on-treatment measurements recorded for serum creatinine and eGFR, including risk differences and 95% CIs (using CMH method). The following change from baseline categories will be analyzed:

- Creatinine (≥1.5 x BL, ≥2.0 x BL)
- eGFR ((≥25% decrease from BL, ≥50% decrease from BL)

In addition, shifts from baseline to each on-study evaluation will be presented for eGFR, based on the standard eGFR categories as defined in Table 7.

**Table 7: Stages of CKD**

| CKD Stage | Glomerular Filtration Rate (mL/min/1.73 m2) |
|---|---|
| G1/Normal | ≥90 |
| G2 | 60 – 89 |
| G3 | 30 – 59 |
| G3a | 45-59 |
| G3b | 30-44 |
| G4/5 | <30 |

CKD = chronic kidney disease; G = Group

Shifts from baseline to Month 18 in urinary albumin to creatinine ratio (mg/g) will also be presented using the categories defined in Table 8.

**Table 8: Cutoffs of Albumin to Creatinine Ratio**

| Category | Albumin to Creatinine Ratio (mg/g) |
|---|---|
| Normal | <30 |
| Microalbuminuria | 30 – 300 |
| Macroalbuminuria | >300 |

## 10.6. Pregnancies

If any subject gets pregnant during the course of the study, a by-subject data listing will be provided including all data on the Pregnancy eCRF.

## 10.7. Vital Signs

The results, change and percentage change from Baseline to each on-study evaluation visit will be summarized for oral temperature, sitting heart rate, and respiratory rate.

All vital sign data will be presented in by-subject data listings.

## 10.8. Electrocardiograms

The central read ECG data will be analyzed based on methodology recommended in the International Conference on Harmonisation (ICH) E14 guideline, The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Nonantiarrhythmic Drugs.

Baseline is defined as the mean of all available evaluations before treatment. Descriptive statistics of ECG parameters (time between 2 consecutive R waves [RR], PR, QRS, QT, and QT interval corrected by the Fridericia's formula [QTcF]) at baseline and at each post-baseline timepoint will be summarized by treatment group; absolute changes from baseline will also be summarized.

A categorical summary of abnormal QTcF values will be presented by treatment group. The number of subjects with values of >450 msec, >480 msec, and >500 msec will be presented, and

the number of subjects with change from baseline values of >30 msec and >60 msec will also be presented.

Overall interpretation results for ECGs and the investigator interpretation results are collected as normal, abnormal not clinically significant (NCS), and abnormal clinically significant (CS). Subjects whose interpretations shift from normal to abnormal (CS or NCS) will be listed separately, including description of the abnormality and any associated comments.

All ECG results will be presented in by-subject data listings.

# 11. CHANGES TO PLANNED ANALYSES

As of this date, the Month 18 IA analyses for resubmission have been conducted per FDA agreement following SAP addendum V2.0 (dated 25April 2022).

The key changes for this SAP are summarized in the table below:

| **Key Changes** | **Rationale** | **Contents** |
|---|---|---|
| Changed to one clinic outcome interim analysis at 66% of information fraction | One IA 66% information fraction will balance preserving type I error with obtaining a reasonable estimate of clinical outcomes efficacy while minimizing the potential for any increases of patient discontinuations as a result of availability of commercial OCA should the agency grant approval of OCA for the NASH indication. | Changed to IA at 66% of information fraction from two IAs at 50% and 80% before end of study analysis under section 2.2.<br>Updated Alpha spending and HR boundaries under section 3.2. |
| Added Estimands for clinical composite outcomes for end of study analyses | Per current regulatory agencies, the estimand approach following the ICH request for the definition of Estimands should be defined in clinical trials and the statistical analysis plan. | Estimands including strategies to handle the intercurrent events are provided under section 2.3.2.1. |
| Updated monitoring plan | The monitoring of clinical outcome events will be used to set up the timeline to initiate and prepare for IA and end of study analysis per FDA's comment. | Updated Monitoring plan for IA and End of study analysis under section 6.4.1. |

| Key Changes | Rationale | Contents |
|---|---|---|
| Included description of consensus method | Consensus Method has been used to re-read the liver biopsies samples for histological endpoint estimates at Month 18 interim analysis resubmission per FDA agreement. It will be used for subsequent histological analysis for clinical outcomes assessment. | Description of Consensus Method under section 6.2.2. |
| Added Safety Objectives to be assessed and additional AESIs | Based on the FDA feedback after Month 18 interim analysis submission according to protocol version 11.0 (Dated 04May2021) | Added additional analyses for AESIs under section 10.4.<br>Added additional analyses based on adjudicated data by Renal, Cardiovascular Diseases and Hepatic Safety Committee under section 10.4 |
| Added Analysis of Covid-19 impact on the study | Following FDA guidance: Conduct of Clinical Trials of Medical Products During the COVID-19 Public Health Emergency | Impact of Covid-19 on protocol deviations under section 5.6<br>Incidence of Covid-19 infection under section 10.3<br>Incidences of AEs before and after Covid-19 under section 10. |
| Updated lab analysis | Per FDA feedbacks, additional lab data analyses related to cardiovascular, renal and hepatic injuries from both local and central laboratories have been produced in Month 18 IA resubmission and will be also done at the end of study | Updated Serum Chemistry Lipids analysis under section 10.4.9.4<br>Updated Liver Injury related lab data analyses under section 10.5.4.<br>Updated Renal injury related lab data analyses under section 10.5.6 |
| [redacted] | | |

# 12. REFERENCES

Guidance for Industry *E9: Statistical Principles for Clinical Trials*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), September 1998.

Guidance for Industry *E9(R1): Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), May 2021.

CKD-EPI Chronic Kidney Disease Epidemiology Collaboration. Implementation of the 2021 CKD-EPI Equations, October 2021.

Guidance for Industry *E14 The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs,* U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), October 2005.

Lan KKG, DeMets DL. Discrete sequential boundries for clinical trials. Biometrika. 1983;70(3):659-3.

Mayo MJ, Parkes J, Adams-Huet B, Combes B, Mills AS, Markin RS, R, Wheeler D, Contos M, West AB, Saldana S, Getachew Y, Butsch R, Luketic V, Peters M, Di Bisceglie A, Bass N, Lake J, Boyer T, Martinez E, Boyer J, Garcia-Tsao G, Barnes D, Rosenberg WM. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. Hepatology 2008; 48: 1549-1557.

Reboussin DM, DeMets DL, Kim KM, Lan KK. Computations for group sequential boundaries using the Lan-DeMets spending function method. Control Clin Trials. 2000; 21:190-207.

Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons; 1987.

Vansteelandt S, Carpenter J, Kenward M. Analysis of Incomplete Data Using Inverse Probability Weighting and Doubly Robust Estimators. Methodology 2010; Vol. 6(1):37–48.

# APPENDIX A. DATA CONVENTIONS AND STANDARD PROGRAMMING

## DATA PRESENTATION CONVENTIONS:

The precision of original measurements will be maintained in summaries, when possible. Minimum and Maximum will be presented with the same decimal places as the raw data. Means, medians, SEMs, and SDs will be presented with an increased level of precision, where means, medians, and percentiles will be presented to one more decimal place than the raw data, and the SEMs and SDs will be presented to two more decimal places than the raw data. In general, the decimal places should not exceed 3 decimal places, unless appropriate.

For tables where rounding is required, rounding will be done to the nearest round-off unit. For example, when rounding to the nearest integer, values ≥XX.5 will be rounded up to XX + 1 (eg, 97.5 will round up to 98), while values <XX.5 will be rounded down to XX (eg, 97.4 will round down to 97).

Percentages based on frequency counts will be based on available data, and denominators will generally exclude missing values. Percentages based on frequency counts will be presented as whole numbers (no decimal places), and values less than 1% will be presented as "<1%." Values less than 100% but that round up from 99.5% to 100% will be presented as ">99%."

Quantitative laboratory tests containing less than (<) and greater than (>) symbols are test results that are below and above quantifiable limits, respectively. In order to retain these values for analysis purpose, the following imputations will be done within the analysis datasets:

For laboratory test results that are below the quantifiable limit:

Imputed laboratory results = (numeric portion of the result) x 0.9.

For laboratory test results that are above the quantifiable limit:

Imputed laboratory results = (numeric portion of the result) x 1.1.

Date variables will be formatted as DDMMMYYYY for presentation. Time will be formatted in military time as HH:MM for presentation.

All output will be incorporated into Microsoft Word rich text format (.rtf) files, sorted, and labeled according to the ICH recommendations, and formatted to the appropriate page size(s).

## PROGRAMMING CONVENTIONS AND STANDARD CALCULATIONS:

### Partial and Missing Dates

If only a partial date is available, and it is required for analysis (for example, duration or time to event analysis), the following general imputation rules will be applied. More details will be provided below for each specific safety endpoints.

- Start date:
  - For missing day only – Day will be imputed as the first day of the month (ie, 1).

- For missing day and month – Day and month will be imputed as the first day of the year (ie, 1 January).
- Missing dates will not be imputed

- Stop date:
  - For missing stop day only - Day will be imputed as the last day of the month (ie, 28, 29, 30, or 31).
  - For missing stop day and month - Day and month will be imputed as the last day of the year (ie, 31DEC).
  - Imputed stop dates must be on or after the start date.

Partial Date Imputation by Safety Endpoints:

- Diagnosis date: Follow the general imputation rules above for start date
- Concomitant medication Start date: Follow the general imputation rules with the following exceptions:
  - If the partial date falls in the same month and year as the first dose date, then the partial date will be imputed to equal the first dose date
  - If the partial date falls in the same year as the first dose date, then the partial date will be imputed to equal the first dose date

- Investigational Product (OCA): Consider only exposure records with non-missing dose.
  - If start date is partial or missing, then start date = non-missing stop date of the preceding exposure record +1 if there is a preceding record (including interruptions and drug holiday). If both dates are partial, choose the date in the middle after imputing the earliest date as the first day in the month and the latest date as the last day in the month.
  - If stop date is partial or missing, then stop date = non-missing start date of the following exposure record -1. If both dates are partial, choose the date in the middle after imputing the earliest date as the first day in the month and the latest date as the last day in the month.
  - Missing stop date in the last exposure record will be imputed by the earliest of
    1. Last exposure date in the EOT/EOS CRF.
    2. End of treatment visit date
    3. Last visit before the end of treatment
    4. End date of study discontinuation, death date or data-cutoff date

- Adverse Events:
  - Missing start day of the first record of AEs with the same system organ class (SOC) and preferred term (PT): Day will be imputed as the first day of the month with the following exception: if the partial date falls in the same month and year as the first dose date, then the partial date will be imputed to equal the first dose date.
  - Missing start day and month of the first record of AEs with the same system organ class (SOC) and preferred term (PT): Day and month will be imputed as the first day of the year with the following exception: if the partial date falls in the same year as the first dose date, then the partial date will be imputed to equal the first dose date
  - Imputed start dates must be prior to the stop date.
  - Missing stop day only: Day will be imputed as the latest of the last day of the month (ie, 28, 29, 30, or 31).
  - For missing stop day and month - Day and month will be imputed as the last day of the year (ie, 31DEC). Imputed stop dates must be on or after the start date.

## Last Contact Date:

Last contact date for time-to-event analysis is calculated as follows:

- Ongoing Subjects (including subjects on-treatment and subjects that discontinued treatment but on-study): DCO date
- Subjects that discontinued study: The date of study discontinuation
- Subjects lost to follow up: The latest of last assessment, last visit, and last dose date

## Standard Calculations:

Variables requiring calculation will be derived using the following formulas:

- Days – A duration expressed in days between one date (*date1*) and another later date (*date2*) will be calculated using the following formulas:

  duration in days = date2 – date1 + 1, where date1 ≥first dose date

  duration in days = date2 – date1, where date1 <first dose date
- Months – A duration expressed in months is calculated as the number of days divided by 28 for study-related duration.
- Years – A duration expressed in years between one date (*date1*) and another date (*date2*) is calculated using the following formulas:

  duration in years = (date2 – date1 + 1)/(# of days in the year), where date1 ≥first dose date

duration in years = (date2 – date1)/ (# of days in the year), where date1<first dose date

# of days in the year= 336 is based on 12 months with 28 days per month for study-related duration calculations.

# of days in the year= 365.25 for age, years since NASH diagnosis, patient-year and other duration calculations that are not study-related

- Age – Age is calculated as the number of years from the date of birth (*DOB*) to the specified date, eg, date of informed consent (*DOIC*). If the month of DOIC <month of DOB or the month of DOIC=DOB and the day of DOIC <day of DOB, then the following formula is used:

  age (years) = year of DOIC – year of DOB – 1.

  Otherwise, the following formula is used:

  age (years) = year of DOIC – year of DOB.

  If only Year is provided in DOB then use July 1 for the month and day

- Date of birth in Year format: To convert the date of birth format from Day/Month/Year to Year only, the following will be applied:

  If birth date before July 1, then use year of birth= reported year -1

  If birth date after July 1, then use year of birth= reported year

- Height – Height entries made in inches (in) are converted to centimeters (cm) using the following formula:

  height (cm) = height (in) × 2.54

- Weight – Weight entries made in pounds (lb) are converted to kilograms (kg) using the following formula:

  weight (kg) = weight (lb) / 2.2046

- Temperature – Temperature entries in degrees Fahrenheit are converted to degrees Celsius using the following formula:

  temp (degrees Celsius) = 5 / 9 × (temp [degrees Fahrenheit] – 32)

- Body Mass Index (BMI) – BMI is calculated using height (cm) and weight (kg) using the following formula:

  BMI (kg/m2) = weight (kg) / ([height (cm)/100]2)

- Change from Baseline – Change from Baseline will be calculated as:

  Change = post Baseline value – Baseline value

- Percentage change from Baseline – Change from Baseline will be calculated as:

  Percentage change from Baseline = ([post Baseline value – Baseline value] / Baseline value) × 100

- NAS (0-8) is derived as the sum of three components of standardized histologic feature scoring system for liver biopsies:

  NAS = steatosis score (0-3) + lobular inflammation score (0-3) + hepatocellular ballooning score (0-2)

- MELD score will be reported as an integer and is derived using the following formula (OPTN):

  Candidates who are at least 12 years old receive an initial MELD(i) score equal to:

  0.957 x Loge(creatinine mg/dL) + 0. 378 x Loge(bilirubin mg/dL) + 1.120 x Loge (INR) + 0.643 Laboratory values less than 1.0 will be set to 1.0 when calculating a candidate's MELD score.

  The maximum MELD score is 40. The MELD score derived from this calculation will be rounded to the tenth decimal place and then multiplied by 10.

  For candidates with an initial MELD score greater than 11, the MELD score is then re-calculated as follows:

  MELD = MELD(i) + 1.32*(137-Na) – [0.033*MELD(i)*(137-Na)]

  Sodium values less than 125 mmol/L will be set to 125, and values greater than 137 mmol/L will be set to 137. MELD Na will be rounded to the nearest integer.

  MELD Score computed from lab results for samples collected in the same visit date. Any exceptional cases, prior to protocol version X, will be documented

- The eGFR will be calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) 2021 equation (CKD-EPI Chronic Kidney Disease Epidemiology Collaboration 2021):

$$eGFR(mL/min/1.73\ m^2) = 142 \times \left(\frac{S_{cr}}{\alpha}\right)^{-0.241} \times 0.993^{age} \times 1.012\ if\ S_{cr} \leq \alpha \text{ for female}$$

  - $= 142 \times (Scr/\alpha)^{-1.200} \times 0.993^{age} \times 1.012\ if$ Scr $> \alpha$ for female
  - $= 142 \times (Scr/\alpha)^{-0.302} \times 0.993^{age}\ if$ Scr $<= \alpha$ for male
  - $= 142 \times (Scr/\alpha)^{-1.200} \times 0.993^{age}\ if$ Scr $> \alpha$ for male

  Where Scr (standardized serum creatinine) = mg/dL, α = 0.7 if female and 0.9 if male, age = years.

[redacted]

## Baseline Definition:

- The table below provides baseline definition for the endpoints to be analyzed. If the endpoint doesn't have assessment except at Day 1 (such as PRO) then baseline will be Day 1 value
- High/Low/Normal will be derived for baseline values, calculated as the mean of all pre-treatment results, based on the ULN and LLN.
- Subjects with repeated baseline: For the purpose of change from baseline, if the subject has a repeated baseline after IP re-challenge, change from baseline will based on the 1st baseline value

**Baseline Definition**

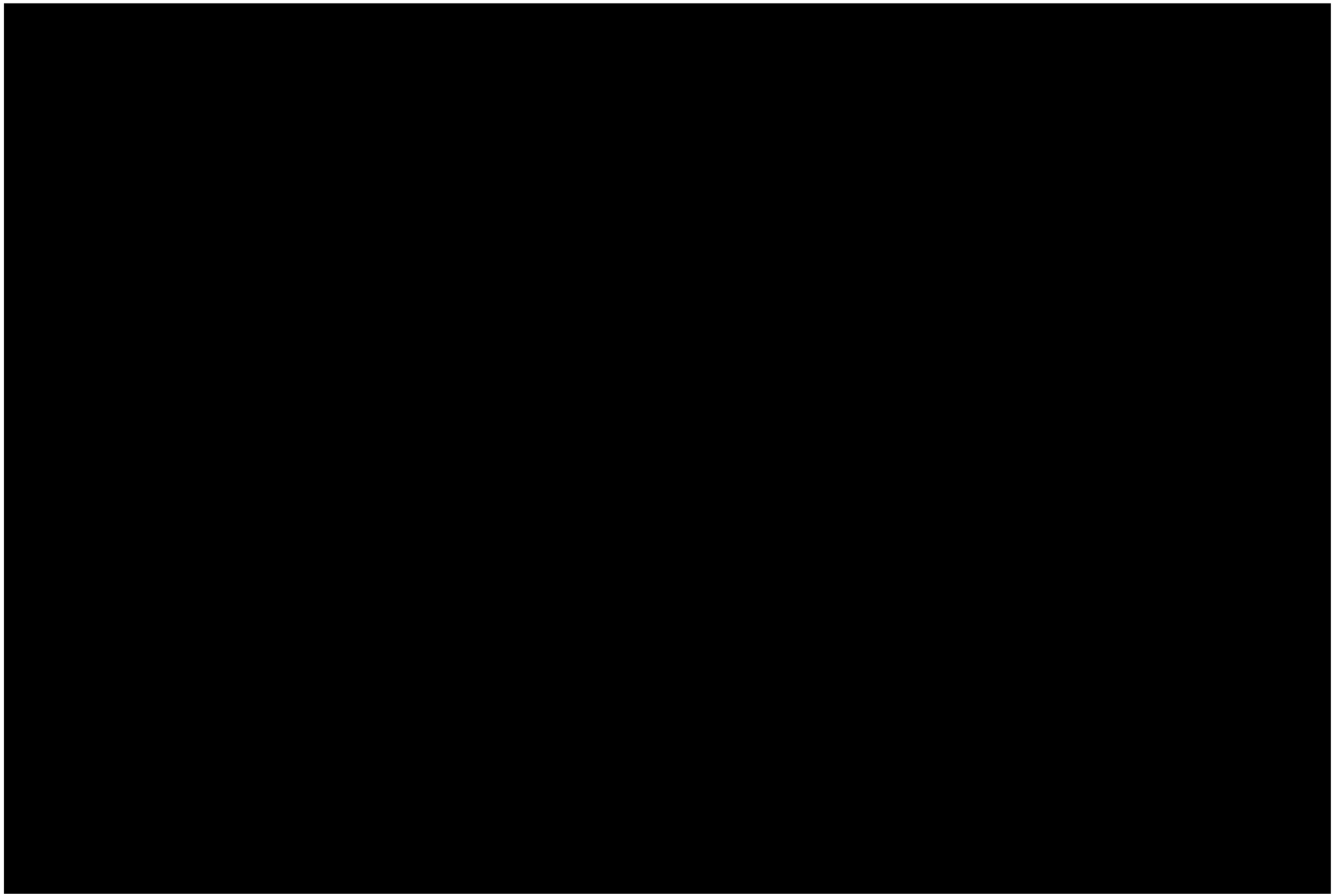
| Assessment | Baseline Definition |
|---|---|
| Serum Chemistry, Hematology and Coagulation | The mean of all available evaluations prior to the first administration of investigational product. |
| Lipoprotein | The last fasted assessment prior to the first administration of investigational product |
| Markers of glucose metabolism | The last fasted assessment prior to the first administration of investigational product |
| Cardiovascular Assessment | The last assessment prior to the first administration of investigational product |
| ECG | The mean of all available evaluations prior to the first administration of investigational product. |
| Anthropometric measures | The mean of all available evaluations prior to the first administration of investigational product |
| Vital signs | The mean of all available evaluations prior to the first administration of investigational product |
| Exploratory Biomarkers | The mean of all available evaluations prior to the first administration of investigational product. |

| | |
|---|---|
| Pharmacodynamic Parameters | The mean of all available evaluations prior to the first administration of investigational product. |

## Additional Conventions:

1. Normal reference ranges may change during the study due to assays re-validated with new reference ranges. For example, reference ranges used in the 747-303 study including INR, APT, APPT, Bicarbonate have been updated during the study.
   - INR, APT, and PPT: Changes in normal reference ranges were implemented to reflect appropriate ranges based on collected data overtime. The update in reference ranges does not reflect changes in sample analysis and consequently the same value corresponding to reference ranges should be considered the same. The latest reference range will be used in summaries. In listings, reference ranges will be reported as received from ICON lab.
2. Concomitant medications:
   - Use of CMSCAT and CMCLASCD versus ATC (level 2, level 4) for hypertensive medications and diabetic medications along with indication.
3. AEs: SMQ
   - Hepatic disorder: Hepatic Disorders SMQ, excluding the following sub-SMQs: congenital, familial, neonatal and genetic disorders of the liver; Liver infections; alcohol related sub SMQ as one of the excluded sub SMQ, and Pregnancy-related hepatic disorders.
   - Cardiac events: the MedDRA Broad SMQs of embolic and thrombotic events or Ischemic heart disease, or Central nervous system vascular disorders narrow SMQ.
   - Renal AEs: Renal Standardized MedDRA query (SMQ): acute renal failure, chronic kidney disease, proteinuria, renovascular disorders, or tubulointerstitial diseases
   - Dyslipidemia: Dyslipidemia SMQ
   - Gallbladder Disease and Related Complications: Gallbladder Related Disorders narrow SMQ and Gallstone Related disorders narrow SMQ, with additional PTs of Biliary abscess, Biliary sepsis, Biliary tract infection, Gallbladder abscess, Gallbladder empyema, Bile duct necrosis, Bile duct obstruction, Bile duct stenosis, Biliary colic, Cholangitis, Cholangitis acute, Cholangitis chronic, Cholecystocholangitis, and Perforation bile duct.
   - Pancreatitis: Acute Pancreatitis narrow SMQ
   - Hyperglycemia: Hyperglycaemia/new onset diabetes mellitus narrow SMQ
4. CTCAE Clarifications:
   - CTCAE for PH is based on blood test and not urine test

- CTCAE for glucose depends on fasting glucose test for grade 1-2 and any available glucose test (fasting or non-fasting) for grade 3-4.
- Lab units in CTCAE grade derivation: Values in SI values will be used to derive CTCAE classification



# APPENDIX C. PREFERRED TERMS TO IDENTIFY POTENTIAL KIDNEY INJURY

**Table C1: Preferred Terms (List 1)**

| |
|---|
| Acute kidney injury |
| Acute phosphate nephropathy |
| Acute prerenal failure |
| Anuria |
| Cardiorenal syndrome |
| Continuous haemodiafiltration |
| Crush syndrome |
| Crystal nephropathy |
| Haemolytic uremic syndrome |
| Hepatorenal failure |
| Nephritis |
| Nephropathy toxic |
| Pancreatorenal syndrome |
| Postoperative renal failure |
| Postrenal failure |
| Prerenal failure |
| Renal failure acute |
| Renal injury |
| Renal ischemia |
| Renal tubular injury |
| Renal tubular necrosis |
| Traumatic anuria |
| Tubulointerstitial nephritis |
| Urate nephropathy |
| Urine output decreased |

**Table C2: Preferred Terms (List 2)**

| |
|---|
| Azotemia |
| Blood creatinine abnormal |
| Blood creatinine increased |
| Blood urea increased |
| Blood urea nitrogen/creatinine ratio increased |
| Creatinine renal clearance abnormal |
| Creatinine renal clearance decreased |
| Dialysis |
| Glomerular filtration rate abnormal |
| Glomerular filtration rate decreased |
| Haemodialysis |
| Haemofiltration |
| Hypercreatinemia |
| Hypercreatininemia |
| Metabolic nephropathy |
| Neonatal anuria |
| Oedema due to renal disease |
| Renal disorder |
| Renal failure |
| Renal failure neonatal |
| Renal function test abnormal |
| Renal impairment |
| Renal insufficiency |
| Renal tubular disorder |
| Renal tubular dysfunction |
| Subacute kidney injury |
| Urea renal clearance decreased |
| Urinary tract toxicity |

# APPENDIX D. LIVER LABORATORY CRITERIA FOR MONITORING FOR POTENTIAL HEPATIC INJURY

| Baseline | Treatment Emergent | Signs and Symptoms of Hepatic Injury[a] | Action |
|---|---|---|---|
| ALT Normal/Near Normal[b] | ALT ≥5x ULN<br>TB normal<br>Gilbert's[c]: no change in BL TB | No | • Repeat laboratory testing including CPK within 2-4 days |
| | ALT ≥3x ULN<br>Regardless of TB or DB | Yes | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| | ALT ≥8x ULN<br>Regardless of TB or DB | No | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| | ALT ≥3x ULNTB ≥2x ULN<br>Gilbert's[c]: DB ≥2x ULN or INR >1.5 | No | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| ALT Elevated[b] | ALT ≥3x BL or ≥300 U/L, whichever occurs first<br>TB normal<br>Gilbert's[c]: no change from BL TB | No | • Repeat laboratory testing including CPK within 2-4 days |
| | ALT ≥2x BL or ≥300 U/L, whichever occurs first<br>Regardless of TB or DB | Yes | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| | ALT ≥5x BL or ≥500 U/L, whichever occurs first<br>Regardless of TB or DB | No | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| | ALT ≥2x BL or ≥300 U/L, whichever occurs first<br>TB ≥2x ULN<br>Gilbert's[c]: DB >2x ULN or INR >1.5 | No | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| TB (>1.5 mg/dL exclusion criterion 747-303) | TB ≥2.0 mg/dL and DB >30% ofTB | Yes/No | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |

| Baseline | Treatment Emergent | Signs and Symptoms of Hepatic Injury[a] | Action |
|---|---|---|---|
| ALP ≤ULN[d] | ALP ≥2x BL and >ULN | No | *If ALP <250 U/L, then:*<br>• Repeat laboratory testing including CPK within 2-4 days<br><br>*If ALP ≥250 U/L, then:*<br>**Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| | | Yes | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| ALP >ULN[d] | ALP ≥2x BL or >300 U/L, whichever occurs first | No | *If ALP <300 U/L, then:*<br>• Repeat laboratory testing including CPK within 2-4 days<br><br>*If ALP ≥300 U/L, then:*<br>**Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |
| | | Yes | **Interrupt**<br>• Repeat laboratory testing including CPK within 2-4 days<br>• Collect PK sample within 7 days |

ALP = alkaline phosphatase; ALT = alanine aminotransferase; BL = baseline; DB = direct bilirubin; CPK = creatine phosphokinase; DILI = drug-induced liver injury; INR = international normalized ratio; NASH = nonalcoholic steatohepatitis; OCA = obeticholic acid; PK = pharmacokinetic; TB = total bilirubin; ULN = upper limit of normal

[a] Signs and symptoms of hepatic injury include severe fatigue, nausea, vomiting, right upper quadrant pain, fever, rash, and eosinophilia.

[b] Baseline ALT is defined as an average of all pre-treatment values from Screening and Day 1. Elevated Baseline defined as ALT ≥1.5x ULN. ALT decreases are common within the first 3-6 months of treatment in NASH subjects receiving OCA. If there is a decrease in ALT for all lab draws through the Month 6 Visit, then a new baseline should be established for subsequent DILI determination, based on the average ALT from the Month 3 and Month 6 Visits.

[c] A subject will be considered to have Gilbert's syndrome if it was established at baseline and documented in the subject's medical history.

[d] Baseline ALP is defined as an average of all pre-treatment values from Screening and Day 1. Modest ALP increases (<50%) from baseline within the first 3 months of treatment are common in NASH subjects receiving OCA and represent an expected pharmacodynamic effect. If there is an increase in ALP through the Month 6 Visit, then a new baseline should be established for subsequent DILI determination, based on the average ALP from the Month 3 and Month 6 Visits.