OBETICHOLIC ACID (OCA)

747-304

A Phase 3, Double-Blind, Randomized, Placebo-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of Obeticholic Acid in Subjects with Compensated Cirrhosis due to Nonalcoholic Steatohepatitis

Statistical Analysis Plan

Author:	Min Mo, Associate Director of Biostatistics
Protocol Version and Date:	Version 7: 16 June 2021
Analysis Plan Version:	Version 3.0
Analysis Plan Date:	15 September 2022

ClinicalTrials.gov Identifier: NCT03439254

SPONSOR'S APPROVAL OF THE STATISTICAL ANALYSIS PLAN

Intercept 747-304 SAP v3 15Sep2022 747-304 Statistical Analysis Plan v3 15Sep2022 (16.1.9)
(Published, v1.0)
VV-CLIN-017745
Product: INT-747 **Document Type:** Clinical › Statistics › Statistical Analysis Plan

[redacted]
Regulatory

Approve
Task: Approve-no signature page
9/15/2022 2:41 PM EDT

[redacted]
Clinical

Approve
Task: Approve-no signature page
9/15/2022 1:41 PM EDT

[redacted]
Clinical

Approve
Task: Approve-no signature page
9/15/2022 1:38 PM EDT

TABLE OF CONTENTS

LIST OF TABLES

[redacted]

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or Specialist Term	Explanation
AE	adverse event
AIM-NASH	AI-based histologic Measurement of NASH
[redacted]	
ANCOVA	analysis of covariance
ApoA-1	apolipoprotein A-1
ApoB	apolipoprotein B
ApoE	apolipoprotein E
[redacted]	
AUDIT	Alcohol Use Disorders Identification Test
BAS	bile acid sequestrants
BMI	body mass index
C4	7α-hydroxy-4-cholesten-3-one
CAC	Cardiovascular Adjudication Committee
CI	confidence interval
[redacted]	
CLDQ	Chronic Liver Disease Questionnaire
CMH	Cochran-Mantel-Haenszel
CRN	Clinical Research Network
CS	clinically significant
CTCAE	Common Terminology Criteria for Adverse Events
CP	Child-Pugh
DB	Double blind
DILI	drug-induced liver injury
ECG	electrocardiogram
EGD	esophagogastroduodenoscopy
eCRF	electronic case report form
EDC	electronic data capture
ELF	Enhanced Liver Fibrosis
EOS	end of study
EOT	end of treatment

Abbreviation or Specialist Term	Explanation
ET	early termination
FDA	Food and Drug Administration
FIB-4	Fibrosis-4
[redacted]	
FLINT	Farnesoid X receptor Ligand OCA in Nonalcoholic Steatohepatitis Treatment
[redacted]	
glyco-OCA	glycine conjugate of obeticholic acid
HbA1c	hemoglobin A1c
HCC	hepatocellular carcinoma
HDL	high-density lipoprotein
HE	hepatic encephalopathy
[redacted]	
INR	international normalized ratio
ITT	Intent-to-Treat
IWRS	interactive web response system
KM	Kaplan-Meier
LDL	low-density lipoprotein
LS	least-squares
MACE	major adverse cardiovascular events
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MELD	model for end-stage liver disease
mITT	modified intent-to-treat
MNAR	missing not at random
NAFLD	nonalcoholic fatty liver disease
NAS	nonalcoholic fatty liver disease activity score
NASH	nonalcoholic steatohepatitis
NCS	not clinically significant
[redacted]	

Abbreviation or Specialist Term	Explanation
OCA	obeticholic acid
OLE	open-label extension
PBC	primary biliary cholangitis
[redacted]	
PCSK9	proprotein convertase subtilisin/kexin type 9
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
PP	per protocol
PT	prothrombin time
PTT	partial thromboplastin time
QTcF	QT interval corrected by the Fridericia's formula
RTSM	Randomization and trial supply management
SAP	statistical analysis plan
SCORE	systemic coronary risk evaluation
SMQ	Standardized MedDRA Query
TE	transient elastography
TEAE	treatment-emergent adverse event
[redacted]	
ULN	upper limit of normal
US	United States
VLDL	very low-density lipoprotein
VAS	visual analog scale
[redacted]	

1. SCOPE

This statistical analysis plan (SAP) provides a detailed description of the strategy, rationale, and statistical techniques to be used to meet the objectives of Study 747-304, which includes a double-blind (DB) phase and an open-label extension (OLE) phase. The SAP provides additional details on the statistical analyses that were described in the protocol. The SAP will be finalized and signed off prior to database lock and study unblinding for the analysis of the DB phase. Any deviations from the methods specified in this SAP will be documented in the clinical study report (CSR). If additional analyses are required to supplement the planned analyses described in this SAP, they will be identified as post-hoc.

The following is a summary of key changes from Version 1 to Version 2:

- Reorganization and ranking of secondary and exploratory objectives. Notable changes are the inclusion of non-invasive liver fibrosis endpoints (previously considered exploratory) as secondary endpoints.
- [redacted] Ishak scoring system and [redacted]
- Modified section on estimand for primary efficacy (Section 7.1.1) by clarifying definition of the intercurrent event and confirming the treatment policy strategy will be applied.
- Included section for handling of missing data, added details for multiple imputation (i.e. covariates to be included in the model and sample code with seed), and as recommended by FDA added tipping point analysis as a sensitivity analysis to assess the impact of missing data on the primary efficacy endpoint.
- The sensitivity analysis for subjects missing biopsies due to COVID-19 (Section 7.6.7 in Version 1.0) was removed. We take the FDA's recommendation of assuming missingness at random. The multiple imputation methodology will cover these scenarios.

Additional changes and clarifications from Version 2 to Version 3 include:

- Secondary clinical outcomes composite endpoint has been expanded to include gastroesophageal varices confirmed by endoscopy (not adjudicated). This has been added to the End of the DB and OLE phases. The presence or absence of varices and the date of occurrence is captured in the eCRF as objective evidence and therefore adjudication is not necessary.
- Day 1 was removed from MMRM change from baseline, since Day 1 is baseline
- Added clarifying note for the derivation of Resolution of NASH objective for End of DB phase
- Analysis for subjects with improvement in fibrosis by 3 stages will no longer be analyzed by histology and AIM-NASH.

2. STUDY OBJECTIVES

2.1. Primary Objectives Assessed at the End of the DB Phase

The primary objective is to evaluate the effects of OCA treatment compared with placebo on:

- Histological improvement in fibrosis by assessing the percentage of subjects with improvement in fibrosis by at least 1 stage with no worsening of nonalcoholic steatohepatitis (NASH) defined as no increase in hepatocellular ballooning or lobular inflammation, using the NASH Clinical Research Network (CRN) scoring system, from Baseline to the end of the Double-Blind Phase

2.2. Secondary Objectives Assessed at the End of the DB Phase

The secondary objectives are to evaluate the effects of OCA treatment compared with placebo on:

- Noninvasive assessments of liver disease assessed by serum markers and imaging tests. These include the following:
  - Assessment of liver stiffness by transient elastography (FibroScan)
  - Calculated laboratory-based score of fibrosis-4 (FIB-4)
  - Blood-based marker of liver fibrosis (enhanced liver fibrosis [ELF])
- Resolution of NASH defined as overall histopathological 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, using the NASH CRN scoring system, from Baseline to the end of the Double-Blind Phase

  Note: This analysis will be derived based on NAFLD activity score (NAS) of 0 for ballooning and 0-1 for inflammation, using the NASH CRN scoring system, from Baseline to the end of the Double-Blind Phase.
- Histological improvement in fibrosis by 2 stages regardless of changes in NAS using the NASH CRN scoring system, from baseline to the end of the Double-Blind Phase
- Occurrence of all-cause mortality and liver-related clinical outcomes for the following events (clinical outcomes composite endpoint):

Adjudicated

- Death (all cause)
- Liver transplant
- Model for End-Stage Liver Disease (MELD) score ≥15
- Worsening of Child-Pugh (CP) score (by at least 2 points)
- Hospitalization (as defined by a stay of ≥24 hours) for:
  - Variceal bleed
  - Hepatic encephalopathy (as defined by a West Haven score of ≥2)

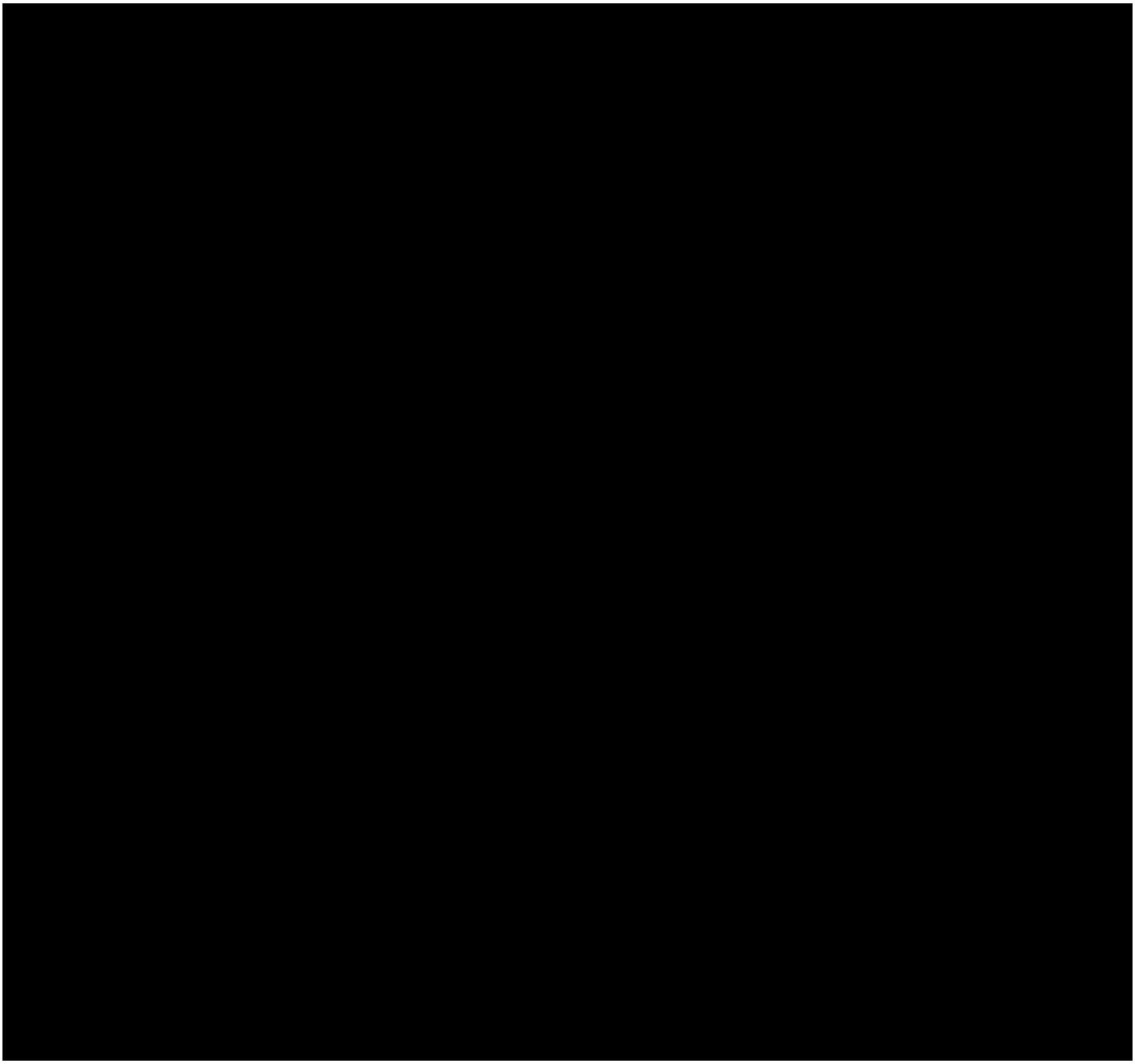
- Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis)

– Ascites secondary to cirrhosis and requiring medical intervention (eg, diuretics or paracentesis)

– Hepatocellular carcinoma (HCC) as confirmed by 2 complementary imaging modalities unless already confirmed by biopsy

Not Adjudicated

– New onset of gastroesophageal varices as confirmed by endoscopy. The occurrence of this event will be used as collected in Esophagogastroduodenoscopy eCRF page.



[redacted]

2.5. Safety Objectives Assessed at the End of the DB Phase

- To evaluate the safety and tolerability of OCA treatment compared to placebo
  - The effect of OCA treatment compared to placebo on the following AEs of special interest (including cardiovascular adverse events, pruritus, renal injury, urolithiasis, gallbladder disease and related complications, pancreatitis, hepatic adverse events, dyslipidemia, and hyperglycemia/ new-onset diabetes mellitus)

- The effect of OCA treatment compared to placebo on the following additional measures and markers:
  - Markers of cardiovascular safety
  - Incidence of adjudicated CV events include core MACE (cardiovascular death, non-fatal MI, non-fatal stroke) and expanded MACE (all causes of death, heart failure (hospitalization and urgent visit), unstable angina (requiring hospitalization), revascularization procedures (coronary or peripheral vascular intervention), transient ischemic attack, arrhythmias, other events related to adverse cardiovascular events).
  - Incidence of adjudicated acute kidney injury events (AKI)
  - Incidence of adjudicated hepatic safety (drug-induced liver injury, DILI) events

2.6. Primary Objective Assessed at the End of the OLE Phase

- To evaluate and summarize the longer-term safety and tolerability of OCA treatment
- To evaluate the effect of OCA on the following noninvasive assessments of liver disease assessed by serum markers and imaging tests:
  - Liver stiffness by transient elastography (FibroScan)
  - Calculated laboratory-based scores of fibrosis-4 (FIB-4)
  - Blood-based markers of liver fibrosis (enhanced liver fibrosis [ELF])
- Noninvasive assessments of liver disease assessed by serum markers and imaging tests
- To summarize the effects of OCA treatment on the occurrence of all-cause mortality and liver-related clinical outcomes for the following events (clinical outcomes composite endpoint):

Adjudicated

  - Death (all cause)
  - Liver transplant
  - MELD score ≥15
  - Worsening of CP score (by at least 2 points)
  - Hospitalization (as defined by a stay of ≥24 hours) for:
    - 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 (e.g., diuretics or paracentesis)
  - HCC as confirmed by 2 complementary imaging modalities or biopsy

Not Adjudicated

- New onset of gastroesophageal varices as confirmed by endoscopy. The occurrence of this event will be used as collected in Esophagogastroduodenoscopy eCRF page.

[redacted]

- Disease progression as assessed by MELD score

3. SUMMARY OF STUDY DESIGN

3.1. Overall Study Design

This Phase 3, DB, randomized, placebo-controlled, multicenter international study will evaluate the efficacy and safety of OCA in subjects with a biopsy-confirmed diagnosis of cirrhosis (based on a fibrosis score of 4 using the NASH CRN scoring system) due to NASH (determined by central reading of liver histology at the time of enrollment in the study). Subjects with hepatic decompensation or CP Class B or Class C cirrhosis are excluded. Subjects who progress to CP Class B or Class C during the study will discontinue investigational product but are expected to be followed through to study closure (or at the discretion of the Sponsor).

During the study, several protocol amendments and revisions were made. The following are the major changes:

- Protocol Version 3 (05 Dec 2017) added the OCA10 mg arm without titration in response to FDA concern about safety of OCA in the PBC program and exposure to 25 mg daily in subjects with cirrhosis. The sample size was increased from 360 to 540.
- Protocol Version 5 (19 Jul 2019) extended the DB phase from 12 months to 18 months, increased the sample size from 540 to 900. [redacted] Uptitration criteria were modified.
- Protocol Version 7 (16 Aug 2020) allowed a 3-month maximum extension of IP if the liver biopsy could not be obtained at M18 due to COVID restrictions. Once obtained, the liver biopsy would still be recorded in the M18 CRFs.

DB Phase (18 Months):

Subjects will be screened for a period of up to 12 weeks before entering the study. Subjects who meet the entry requirements will be randomized in a 1:1:1 ratio to one of the three treatment arms: once daily dosing of OCA 10 mg, OCA 10 mg→ 25 mg (ie, OCA 10 mg with uptitration to OCA 25 mg at Month 3), or placebo, in conjunction with local standard of care. Uptitration will be determined based on the laboratory criteria and safety and tolerability assessments completed prior to Month 3. If uptitration at Month 3 is not feasible, the window for uptitration

may be extended by up to one calendar month, after consultation and agreement with the Medical Monitor. The uptitration review process is detailed in the study-specific Medical Management Plan.

Randomized subjects will be stratified by the presence of type 2 diabetes at enrollment (yes or no). Investigational product (ie, OCA or placebo) will be taken orally, with water, once daily. Efficacy, safety, and laboratory assessments will be evaluated at clinical visits on Day 1, monthly for the initial 6 months (Month 1 through Month 6 Visits), Month 9, Month 12, Month 15, and Month 18.

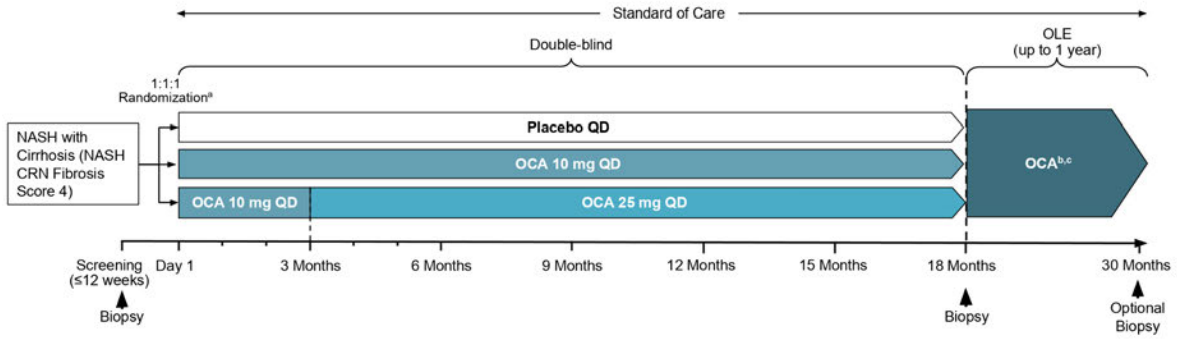
OLE (up to 12 Months):

Subjects who complete the DB Month 12 (prior to version 5) or Month 18 Visit (and continue to receive investigational product) are eligible to enroll into the OLE. Subjects with endoscopic evidence of varices at the end of treatment visit for the double-blind period are not eligible to participate in the OLE Phase. All subjects will receive OCA upon entry into the OLE. Subjects randomized to placebo in the DB phase will be re-randomized to either OCA 10 mg or OCA 10 mg→ 25 mg titration (ie, OCA 10 mg for the first 3 months of the OLE prior to uptitrating to OCA 25 mg at OLE Month 3). Uptitration will be determined based on the same criteria and assessments as employed in the DB phase. Subjects randomized to OCA (10 mg or OCA 10 mg→ 25 mg titration dose) during the DB phase will continue the same dosing regimen they received at the end of the DB-Phase; however, they will undergo dummy titration to maintain study blind until all subjects complete the DB phase and the database is locked.

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

3.2. Study Design Diagram

Figure 1: Study Design Schematic



CRN = clinical research network; NASH = nonalcoholic steatohepatitis; OCA = obeticholic acid; OLE = open-label extension; QD = once daily.

[a] Subjects who meet the entry requirements will be randomized in a 1:1:1 ratio to one of the three treatment arms: once daily dosing of OCA 10 mg, OCA 10 mg → 25 mg (i.e., OCA 10 mg with uptitration to OCA 25 mg at Month 3), or placebo, in conjunction with the standard of care

[b] All subjects will receive OCA upon entry into the OLE: Subjects who received placebo during the Double-Blind Phase will be re randomized 1:1 to either OCA 10 mg or OCA 10 mg → 25 mg (i.e., OCA 10 mg with uptitration to OCA 25 mg at OLE Month 3). Subjects who received OCA during the Double-Blind Phase will continue the same dosing regimen they received at the end of the Double-Blind Phase.

[c] During the OLE period, subjects will return for site visits at Months 1, 2, 3, 4, 5, 6, 9, and 12.

Note:

- Subjects with cirrhosis (based on a NASH CRN fibrosis score 4) due to NASH (determined by central reading of liver histology) will be enrolled in the study. Subjects with hepatic decompensation or CP Class B or CP Class C cirrhosis are excluded.
- Two screening visit assessments will be performed. Screening Visit 1 will occur no more than 12 weeks prior to Day 1, and Screening Visit 2 will occur at least 4 weeks after Screening Visit 1.
- Randomized subjects will be stratified by the presence of type 2 diabetes at enrollment (yes or no).
- The study will remain blinded until all subjects complete the Double-Blind Phase, and the database is locked. To maintain blinding, all investigational product (placebo and OCA) tablets and bottles will be identical.

3.3. Randomization Methodology

This study will be conducted in a double-blind, placebo-controlled manner in DB phase. Enrolled subjects will be randomized in a 1:1:1 ratio to one of the three treatment arms: once daily dosing of OCA 10 mg, OCA 10 mg → 25 mg (ie, OCA 10 mg with uptitration to OCA 25 mg at Month 3), or placebo. Randomization will be done using an RTSM and will be based on a predefined randomization code generated by the Sponsor or designee. Subjects randomized to OCA 10 mg → 25 mg will initiate dosing at OCA 10 mg for 3 months prior to up titrating to OCA 25 mg. Uptitration will be determined based on the laboratory criteria and safety and tolerability assessments completed prior to Month 3.

Randomization of subjects will be stratified by the presence of type 2 diabetes at enrollment (yes or no).

All subjects who enroll into the OLE will receive OCA. Subjects randomized to placebo in the DB phase will be re-randomized to OCA 10 mg or OCA 10 mg → 25 mg arms, with uptitration at OLE Month 3 in the OCA 10 mg → 25 mg arm.

The RTSM will also serve as the investigational product inventory and management system.

The Investigator or designee will be required to register the subject in the RTSM and may be prompted to provide patient data necessary to properly randomize and allocate the subject to treatment. A randomization number will be assigned and investigational product (OCA or placebo) dispensing information will be provided (eg, bottle numbers allocated) while maintaining the study blind.

3.4. Sample Size Determination

During the study, the following protocol amendments and revisions were made that affected sample size. The sample size was increased from 360 to 900 as follows:

- Protocol Version 3 (05 Dec 2017) added OCA10 mg arm without titration. Sample size was increased from 360 to 540.
- Protocol Version 5 (19 Jul 2019) extended DB phase from 12 months to 18 months, increased the sample size from 540 to 900.

A sample size of 300 subjects per group provides at least 90% power to detect a statistically significant treatment difference of 10% between OCA 10 mg → 25 mg titration treatment and placebo groups based on a Chi-square test with 2-sided type I error at 0.05 level, assuming a responder rate of 10% in the placebo group. As there was no previous study performed using the same endpoint in this disease, the assumption was determined based on data from literature, and results from the FLINT study and Month 18 Interim Analysis from Study 747-303.

3.5. Biopsy Reads

Eligibility Read for Screening Visits: During Screening Visits 1 or 2, biopsy data may be collected either from historical unstained slides or from a new procedure if suitable historical slides are unavailable. These are assessed for adequacy and if the subject meets other study eligibility criteria, the slides are sent to a central reader to determine whether the subject meets the inclusion/exclusion criteria with respect to liver histopathology. The screening biopsy was evaluated for fibrosis stage with presence of NASH or not but does not include NAS scores. The SAP refers to the biopsy read for study eligibility as the "Eligibility Read."

Consensus Read for Screening Visits: All slides from screening visits will be re-read by a panel of expert pathologists using a consensus read methodology. Based on communications with the FDA, an updated independent consensus approach for evaluating biopsies was developed using a panel composed of 3 pathologists with experience in NASH histopathology. Slides from all subjects collected during the study will be read using whole slide images (WSIs) employing the independent consensus read approach. The details of the consensus reading process can be found in the Central Histology Manual.

Both fibrosis stage and NAS scores will be available from the Consensus Read. The SAP refers to this biopsy data as the "Consensus Read" for Screening Visits.

Post-Baseline Read for M12/M18/EOT/EOS Visits: A biopsy is required for all subjects at Month 12 or Month 18, or for subjects who discontinue treatment or terminate the study earlier during the DB phase. At Month 18/Month 12 or early termination, biopsy slides will be read in a blinded fashion. The SAP refers to this as the "Post-Baseline Read".

Ishak Criteria: in addition to primary endpoint analysis using the NASH CRN scoring system, biopsy samples will also be scored based on modified Ishak criteria (Ishak 1995) for fibrosis. The Ishak system scores fibrosis using seven categories (0–6), with a score of 5 representing incomplete cirrhosis and 6 representing established cirrhosis.

[redacted]

4. ANALYSIS POPULATIONS AND APPROACHES TO ANALYSIS

4.1. Intent-to-Treat Population

The Intent-to-Treat (ITT) Population will include all randomized subjects and will represent the primary efficacy analyses population. Treatment assignment will be based on the randomized treatment.

4.2. Modified Intent-to-Treat Population

The Modified ITT Population (mITT) will include all ITT subjects except those who are not eligible to dose titrate due to safety or tolerability reasons. Treatment assignment will be based on the randomized treatment.

4.3. Per Protocol Population

The Per Protocol (PP) Population will include all mITT subjects who do not have any major protocol violations that potentially affect the efficacy conclusions. Treatment assignment will be based on the randomized treatment.

4.4. Safety Population

The Safety Population will include all randomized subjects who receive at least 1 dose of investigational product (OCA or placebo). The Safety Population will be the primary population used for safety analyses. Treatment assignment will be based on the actual treatment received. Once subjects are exposed to OCA 25 mg, they will be included in the OCA 25 mg group, even if they get down titrated to 10 mg at some later timepoint.

4.5. PK Population

The PK Population will consist of all subjects who receive OCA, have at least one confirmed analyzable sample, and have no major protocol deviations that may potentially affect the PK analysis. The final determination of subjects included in the PK Population and the samples included in the analysis will be determined based on a review of data before database lock. The PK population will be the population used for OCA PK and PK/PD (ie, exposure-response)

analyses. For the exposure-response analyses, the placebo group from the ITT population will also be included as the reference group representing the absence of drug exposure.

5. GENERAL CONSIDERATIONS

Individual subject data obtained from electronic case report forms (eCRFs), central laboratories, other vendors (eg, eRT), and any derived data will be presented in data listings by subject. Due to COVID-19, some of laboratory assessments are from local laboratories. The local laboratory results will not be summarized, these will be listed in the lab listings, used for safety monitoring.

All summaries will be based on observed values. 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.

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 version 23.0). Concomitant medications will be coded using World Health Organization (WHO) Drug Dictionary (WHODrug Global B3).

5.1. Baseline Definitions

In general, baseline is defined as the last value collected prior to the first administration of the investigational product (IP).

The following lipid parameters and markers of glucose metabolism will be based on the last fasted evaluation prior to the first administration of IP as Baseline: C-peptide, Insulin, Cholesterol, Glucose, HDL Cholesterol, LDL Cholesterol, Triglycerides, LDL Cholesterol, and Glucose.

For selected lab tests (Appendix C), Baseline is defined as the mean of all available evaluations prior to first administration of IP.

5.2. Analysis Windows

For DB phase, analysis windows will be applied on selective assessments. For OLE phase, no visit windows will be applied. Analysis for OLE phase will be done by scheduled visits.

Analysis 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).

Each timepoint window will have the target date as the center and the lower bound is the midpoint between the target date of this window and the preceding one. The upper bound is the midpoint 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 has the same distance from the target date, then the latest assessment will be used, where “latest” will be based on sorting by date, time, record number/sequence number. Table 1 shows the analysis windows of each timepoint.

For post baseline biopsy assessments, there is no analysis window defined. The biopsy at the end of the DB phase (End of DB Phase) will include subjects who had a biopsy at Month 12, and subjects who had a biopsy at Month 18.

Table 1: Analysis Window

	[redacted] **Coagulation,** [redacted] **MELD Score, FIB-4,** [redacted]	[redacted]	**ELF,** [redacted]	**Cardiovas-cular Risk Scores and ECG**	**Anthropometric measures, Liver Stiffness**
DB Phase	**Target Study Day and Window**				
Screening					
Day 1	Day 1	Day 1	Day 1	Day 1	Day 1
Month 1	30 (2, 61)	30 (2, 61)			
Month 3	91 (62, 137)		91 (62, 137)		
Month 6	183 (138, 228)	183 (138, 228)	183 (138, 228)		183 (138, 228)
Month 9	274 (229, 320)				
Month 12	365 (321, 411)	365 (321, 411)	365 (321, 411)	365 (321, 411)	365 (321, 411)
Month 15	457 (412, 502)				
Month 18	548 (503, before OLE)	548 (503, before OLE)	548 (503, before OLE)	548 (503, before OLE)	548 (503, before OLE)
End of DB Phase	Latest post-baseline before OLE	Latest post-baseline before OLE		Latest post-baseline before OLE	Latest post-baseline before OLE

6. SUPPORTING STUDY INFORMATION

6.1. Subject Disposition

Subject disposition will be tabulated by treatment group and all treatment group combined and will include the number of subjects:

DB Phase

- Screened
- Reason for screen failure
- Randomized
- Randomized but not treated
- Uptitrated (dummy or actual) at Month 3/Month 4 during DB phase
- Down titrated (dummy or actual) during DB phase
- Biopsy Collected
  - Performed biopsy at Month 12 but not early termination of investigational product

- Performed biopsy at Month 18 but not early termination of investigational product
- Early termination of investigational product and performed biopsy at early investigational product termination date
- Early termination of investigational product and performed biopsy at Month 12
  - Off IP prior biopsy >90 days
- Early termination of investigational product and performed biopsy at Month 18
  - Off investigational product prior biopsy >90 days
- Early termination of DB phase and performed biopsy at early termination date (optional biopsy)

- Biopsy Not Collected
  - Early terminated investigational product and no biopsy collected
  - Early terminated DB phase and no biopsy collected
  - Completed investigational product and no biopsy collected
- Primary reasons for investigational product discontinuation
- Contacted for Follow-up if Discontinued Investigational Product
- Study Discontinuation prior to end of DB phase
- Primary Reason for DB phase discontinuation

OLE

- Subjects entered OLE phase
- Uptitrated (dummy or actual) at Month 3/Month 4 of OLE phase
- Down titrated (dummy or actual) during OLE phase
- Completed OLE
  - Performed biopsy at the end of OLE
  - No biopsy
- Investigational product discontinuation prior to end of OLE
  - Performed biopsy at early termination visit
  - Performed biopsy at the end of OLE
  - No biopsy
- Primary reasons for investigational product discontinuation
- Contacted for follow-up if discontinued investigational product
- Study discontinuation prior to end of OLE phase
- Primary reason for OLE phase discontinuation

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.

6.2. Demographics and Baseline Disease Characteristics

Demographic variables will include the following:

- Age at the time of informed consent and group (<65 years, ≥65 years, ≥75 years)
- Gender
- Race
- Ethnicity
- Geographic region: North America, Europe, and Rest of World (ROW)

Other baseline characteristics will include the following:

- Hepatocellular ballooning score
- Lobular Inflammation score
- Steatosis score
- Nonalcoholic fatty liver disease activity score (NAS; <6, ≥6)
- Weight (kg)
- Height (cm)
- Body mass index (BMI; kg/m2): Categorical (<30 kg/m^2, ≥30 kg/m^2 - 35 kg/m^2, ≥35 kg/m^2)
- Diabetes status at enrollment (Yes/No)
- [redacted]
- [redacted]
- 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])
- History of cholecystectomy (Yes/No)
- Baseline Alcohol Consumption as assessed by Alcohol Use Disorders Identification Test (AUDIT) Questionnaire (AUDIT Score ≤7, >7)
- Baseline Smoking Status (Yes/No)
- [redacted]: FIB-4, [redacted], ELF, and [redacted]

- Baseline MELD Score
- Baseline Child-Pugh Score

Liver biochemistries, and eGFR results will be summarized as continuous variables and as frequencies for categorizations.

- eGFR (chronic kidney disease epidemiology collaboration [CKD-EPI] calculation) and groups (G1/Normal, G2, G3)
- [redacted]

Demographics and Baseline characteristics will be summarized and presented by treatment group and overall, for each analysis population.

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.

6.3. Baseline Cirrhosis/NASH Disease Characteristics

Baseline Cirrhosis/NASH disease characteristics will be summarized using data collected from the Diagnosis of Cirrhosis/NASH eCRF. Variables include the following:

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

Baseline Cirrhosis/NASH disease characteristics will be summarized and presented by treatment group and overall, for each analysis population. 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 Cirrhosis/NASH disease characteristics will be presented in by-subject data listings.

6.4. Medical History

Verbatim terms will be mapped to Preferred Term (PT) and System Organ Class (SOC) using Medical Dictionary for Regulatory Activities (MedDRA) version 23.0.

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 system organ class.

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

- Cardiovascular
- Pruritus

- Hepatic adverse events
- Dyslipidemia
- Gallbladder disease and related complications
- Pancreatitis
- Renal injury
- Urolithiases
- Hyperglycemia/New Onset Diabetes mellitus

Medical history, including statin use history, will be presented in a by-subject data listing.

6.5. Pretreatment, Prior and New Concomitant Medication

Verbatim terms on eCRFs will be mapped to Anatomical Therapeutic Chemical (ATC) class and Preferred Name using the WHO Drug Dictionary version WHODrug-Global-B3.

Pretreatment concomitant medications are those medications with start and stop prior to the initial dose of investigational product.

Prior concomitant medications are those medication started prior to and continued after the initial dose of investigational product. The following prior concomitant medication will be presented in a separate table:

- Anticoagulants/antiplatelets
- Antidiabetic medications
- GLP-1 Agonists (indication for T2DM or weight loss), includes semaglutide, liraglutide, dulaglutide, bydurean, lixisenatide, exenatide
- Bile acid sequestrants (BAS), also known as BABA-Bile acid binding agents
- Lipid lowering medications
- Statins only
- Methotrexate
- Vitamin E

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.

Pretreatment, 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. A listing will also be provided.

6.6. Medical and Surgical Procedures

Medical and surgical procedures will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.0. Frequency summaries will be provided by SOC, and PT. Prior procedures are defined as those started before the first dose of IP (whether or not ended before the first dose of IP). A concomitant procedure is defined as non-study procedure started during the phase, or non-study procedure started before the phase and ended or remained ongoing during the phase.

6.7. Protocol Deviations

Eligibility deviations specified in the protocol and any other deviation identified during the course of the study that impact analysis of the primary endpoint will be considered a major deviation and subjects meeting any of these deviations will not be included in the PP Population. Major deviations that exclude a subject from inclusion in the PP population will be determined prior to unblinding.

Specifically, Inclusion Criteria #2 will be based on Eligibility Read for Screening Visits. Subjects will NOT be considered as a protocol deviation if Consensus Read for the Screening Visit is not a fibrosis score of 4. All major protocol deviations will be summarized and listed by treatment group.

7. EFFICACY ANALYSES ASSESSED AT THE END OF THE DB PHASE

7.1. Primary Efficacy Endpoint

7.1.1. Estimand

The following table identifies the elements of the intercurrent events and strategies to address for the primary efficacy endpoint.

Table 2: Primary Efficacy Estimand

Estimand		
Target study population	Adult subjects with compensated cirrhosis due to nonalcoholic steatohepatitis	
Primary analysis population	Intent to treat (ITT)	
Endpoint (variable)	Improvement in fibrosis by at least 1 stage with no worsening of NASH (No worsening of NASH is defined as no increase in hepatocellular ballooning or lobular inflammation using NASH CRN scoring system) from Baseline to the end of the DB phase. Biopsy scores are based on the "Consensus Read" at both baseline and post-baseline.	
Summary measure	Percentage of subjects with fibrosis improvement by at least 1 stage and no worsening of NASH from Baseline to the end of the DB phase, including all biopsies collected after treatment discontinuation.	
Analysis	Comparison between placebo vs OCA 10 mg → 25 mg, then placebo vs OCA 10 mg will be performed using a CMH test stratified by the randomization strata (type 2 diabetes at enrollment [yes/no]).	
Intercurrent events/ Strategies to address	Discontinuous treatment due to AE/Death	Treatment policy: For subjects with early discontinuation due to AE or death, the last available post-baseline biopsy will be used.
	Shorter treatment exposure: (1) subjects who discontinue IP prior to termination of the study and continue to follow the regular visit schedule through to study closure (2) subjects who restart IP after a prolonged interruption (ie, longer than 1 month [+2 weeks])	Treatment policy: Treatment effect regardless of intercurrent event.

7.1.2. Hypotheses Testing

The primary efficacy hypothesis of OCA 10 mg → 25 mg titration based on fibrosis improvement without worsening of NASH:

- Null: The percentage of subjects with fibrosis improvement by at least 1 stage and no worsening of NASH, the latter defined as no increase in hepatocellular ballooning or lobular inflammation using NASH CRN scoring system, from Baseline to the end of the DB phase, is equal between placebo and OCA 10 mg → 25 mg titration.
- Alternative: The percentage of subjects with fibrosis improvement by at least 1 stage and no worsening of NASH, the latter defined as no increase in hepatocellular ballooning or lobular inflammation using NASH CRN scoring system, from Baseline to the end of the DB phase, is different between placebo and OCA 10 mg → 25 mg titration.

Note: OCA 10 mg → 25 mg titration includes all subjects randomized to the OCA 10 mg → 25 mg arm regardless of whether they were up-titrated or not.

The primary efficacy hypothesis of OCA 10 mg without titration based on fibrosis improvement without worsening of NASH:

- Null: The percentage of subjects with fibrosis improvement by at least 1 stage and no worsening of NASH, the latter defined as no increase in hepatocellular ballooning or lobular inflammation using NASH CRN scoring system, from Baseline to the end of the DB phase, is equal between placebo and OCA 10 mg.
- Alternative: The percentage of subjects with fibrosis improvement by at least 1 stage and no worsening of NASH, the latter defined as no increase in hepatocellular ballooning or lobular inflammation using NASH CRN scoring system, from Baseline to the end of the DB phase, is different between placebo and OCA 10 mg.

7.1.3. Analyses of Primary Endpoint

A Cochran-Mantel-Haenszel (CMH) test will be performed and stratified by the randomization strata (type 2 diabetes at enrollment [yes/no]). The difference in response rate and associated 95% CI, and p-values will be reported. The response ratio may be additionally presented.

Type I error for the primary efficacy analysis will be controlled at the 0.05 level, adjusting for multiplicity as described in Section 7.6.5.

7.1.4. Handling Missing Data and Early Discontinuation

7.1.4.1. Non-responder Imputation

For the primary analyses, any subject who discontinued treatment due to an AE or died and who does not have an evaluable post-baseline biopsy assessment will be considered a non-responder. That is, for the percentage of responders, the subject will be included in the denominator and will not contribute to the numerator.

7.1.4.2. Multiple Imputation (MI)

Multiple imputation (MI) will be performed to handle other missing data in the primary analysis.

If a subject who does not have post-baseline biopsy assessment and who also discontinued treatment early due to an AE or death, the non-responder imputation will be applied first. Then, for all other missing data scenarios, multiple imputation (MI) methods will be used assuming it is missing at random (MAR).

1000 imputed data sets will be drawn separately for each randomized group, replacing missing outcome values with simulated values using predictive mean matching from a set of imputation models containing all potential prognostic baseline covariates as well as efficacy endpoints. The weighted average method will be fitted to each of the 1000 imputed data sets and the results will be combined using Rubin's rules (Rubin, 1987).

Sample SAS codes are provided in Appendix B.

7.1.4.3. Tipping Point Analysis

A tipping point analysis will be performed to examine the impact of missing data under missing not at random (MNAR) assumption. The tipping point analysis will only be performed for the primary efficacy endpoints when the MI analysis result is statistically significant.

Except for subjects who discontinued due to an AE or died, whose missing data is imputed as non-responder, missing data is first replaced using the MI procedure described in the main analysis. The imputed post-baseline values in the imputed datasets are then subsequently shifted toward a null hypothesis, until the significance of model estimates is overturned. The plausibility of the magnitude in the shift parameter required to overturn the significance is then interpreted from a clinical perspective.

The tipping points that alter the statistical conclusion will be provided.

7.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

- Change from baseline to Month 18 in liver stiffness measurement (LSM) as measured by vibration controlled transient elastography (FibroScan®), absolute (kPa) and percent change
- Change from baseline to Month 18 in FIB-4 (Fibrosis-4)
- Change from baseline to Month 18 in ELF (Enhanced Liver Fibrosis)

If the null hypothesis concerning the primary efficacy endpoint is statistically significant, statistical tests for the secondary endpoints will be performed using a hierarchical strategy. The sequential closed testing gate-keeping procedure is described in Section 7.6.5. Lower level endpoints will be tested only if upper level endpoints are statistically significant at the pre-specified alpha level. OCA 10 mg vs PBO will be tested only if the corresponding comparison for OCA 10-25 mg vs PBO is statistically significant.

7.2.1. Change from Baseline to Month 18 in Liver Stiffness as Measured by Fibroscan® TE Device

Noninvasive radiological methods to assess liver stiffness will be conducted at selected study sites where the respective devices are available. These assessments are taken by vibration controlled transient elastography (TE) method using FibroScan®.

The hypotheses for the OCA 10 mg $\rightarrow$ 25 mg titration group are:

- Null: The change from baseline to Month 18 in liver stiffness median is equal between the placebo and OCA 10 mg $\rightarrow$ 25 mg titration arms.
- Alternative: the change from baseline to Month 18 in liver stiffness median is different between the placebo and OCA 10 mg $\rightarrow$ 25 mg titration arms.

The hypotheses for the OCA 10 mg group are:

- Null: The change from baseline to Month 18 in liver stiffness median is equal between the placebo and OCA 10 mg arms.
- Alternative: the change from baseline to Month 18 in liver stiffness median is different between the placebo and OCA 10 mg arms.

A Mixed model repeated measures (MMRM) analysis will be used to track the trajectory over time in the change from baseline to Month 18 in liver stiffness median. The model will include terms for randomized treatment group, visits (Month 6, Month 12, and Month 18), stratification factor, randomized treatment group by visit interaction and baseline score as covariates. Subject will be included as a random effect and an unstructured covariance matrix can be used assuming convergence can be attained. The principal comparison is at Month 18. The corresponding least square means (lsmeans), the associated standard errors, the difference in lsmeans the associated 95%CI, and p-values will be presented.

7.2.2. Change from Baseline to Month 18 in FIB-4

FIB-4 is a noninvasive assessment of liver disease assessed by a combination of age, ALT and platelet results. FIB-4 will be analyzed using the similar MMRM model as described in Section 7.2.1 at visits Month 1, Month 3, Month 6, Month 12, Month 15, and Month 18.

7.2.3. Change from Baseline to Month 18 in ELF (Enhanced Liver Fibrosis)

ELF is a noninvasive panel of circulating fibrosis markers calculated from serum biomarkers. ELF will be analyzed using the similar MMRM model as described in Section 7.2.1 at visits Month 3, Month 6, Month 12, and Month 18.

7.3. Other Secondary Efficacy Endpoints

Other secondary efficacy endpoints by assessments are as follows:

- Resolution of NASH defined as overall histopathological 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, using the NASH CRN scoring system, from Baseline to the end of the Double-Blind Phase (see Section 7.3.1 for derivation clarification).Histological improvement in fibrosis by 2 stages regardless of changes in NAS using the NASH CRN scoring system, from baseline to the end of the Double-Blind Phase
- Clinical outcome composite endpoint regarding occurrence of all-cause mortality and liver-related outcomes.

The analyses for these secondary efficacy endpoints will be supportive and interpreted descriptively. No adjustments for multiplicity will be made. Both 95% CIs and nominal p-values will be provided.

7.3.1. Percentage of Subjects with Resolution of NASH

Resolution of NASH is defined as no fatty liver disease or fatty liver disease (simple or isolated steatosis) without steatohepatitis AND a NAS of 0 for ballooning and 0-1 for inflammation, using the NASH CRN scoring system, from Baseline to the end of the DB phase.

This endpoint will be derived based on NAFLD activity score (NAS) of 0 for ballooning and 0-1 for inflammation, using the NASH CRN scoring system, from Baseline to the end of the DB Phase.

Percentage of subjects with NASH resolution at Month 12/Month 18 will be analyzed in the same manner to the primary analysis of the primary endpoint described in Section 7.1.3 and Section 7.1.4.1. The CMH test will be performed and non-responder imputation will be used.

7.3.2. Percentage of Subjects with Improvement of Fibrosis by at Least 2 Stages

The percentage of subjects with improvement in fibrosis by at least 2 stages regardless of changes in NAS using the NASH CRN scoring system, from baseline to the end of the DB phase, will be analyzed in the same manner to the primary analysis of the primary endpoint described in Section 7.1.3 and Section 7.1.4.1. The CMH test will be performed and non-responder imputation will be used.

7.3.3. Clinical Outcome Composite Endpoint

Table 3 lists the clinical outcome composite endpoint, as well as the individual components of the composite endpoint. Only adjudicated events will be included in analyses. The time to first occurrence of the clinical outcome composite component and each? individual component endpoints will evaluate the effect of OCA (10 mg and 10 mg $\rightarrow$ 25 mg titration) compared to placebo for the ITT Population.

For the analysis of time to event, a log rank test stratified by the randomization stratification factor will be used. All summaries of incidence will include the associated exact binomial 95% CI. The hazard ratio and its 95% CI will be determined based on a Cox regression model stratified by randomization strata to estimate the magnitude of effect. Subjects who do not experience an event will be censored at the date of last contact. The detail is provided in Section 7.6.1.

In addition, the treatment groups will be compared by the percentage of subjects who reported these adjudicated events.

Table 3: Clinical Outcome Component Endpoints

Clinical Outcomes Composite Endpoint
Occurrence of all-cause mortality and liver-related clinical outcomes for the following events: **Adjudicated** – Death (all cause) – Liver transplant – MELD score ≥15 – Worsening of CP score (by at least 2 points) – Hospitalization (as defined by a stay of ≥24 hours) for: o Variceal bleed o Hepatic encephalopathy (as defined by a West Haven score of ≥2) o Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis) – Ascites secondary to cirrhosis and requiring medical intervention (eg, diuretics or paracentesis) – HCC (as confirmed by 2 complementary imaging modalities or biopsy **Not Adjudicated** – New onset of gastroesophageal varices as confirmed by endoscopy. The occurrence of this event will be used as collected in Esophagogastroduodenoscopy eCRF page

7.4. Exploratory Efficacy Endpoints

[redacted]

[redacted]

7.4.2. Exploratory Efficacy Endpoints by Other Assessments

Table 5 describes the assessments by serum markers, imaging tests, health utilities, and other scoring systems. Descriptive summaries will be provided by treatment group and visit.

7.5. Interim Analyses

This study has no plan for an interim Analyses.

7.6. Statistical Analysis Methods/Issues

7.6.1. General Statistical Methods

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 analysis population unless otherwise specified.

In response analysis [yes/no] of histological endpoints, the comparison between treatment groups will be performed using a CMH test stratified by the randomization strata (type 2 diabetes at enrollment [yes/no]).

For continuous variables, the number of subjects, mean, standard deviation (SD), standard error of the mean (SEM), median, 25th and 75th 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 over time will be analyzed using MMRM, 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 25th, 50th (median), and 75th 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 without any documentation of events will be censored at the date of last contact if they do not participate in OLE, otherwise, they will be censored at the date of last contact or the date of randomization into OLE, whichever occurs first.

7.6.2. Supportive Analyses

7.6.2.1. Analyses on mITT and PP Population

The primary and key secondary endpoint analyses will be conducted in the following populations defined in Section 4, in addition to ITT:

- mITT population
- Per-protocol population

7.6.2.2. Alternative Definition of No Worsening of NASH

In the protocol, the definition of "no worsening of NASH" was defined as no increase in hepatocellular ballooning or lobular inflammation using the NASH CRN scoring system. A supportive analysis will be conducted based on the modified definition of no worsening of NASH as follows:

- Improvement in fibrosis by at least 1 stage with no worsening of NASH defined as no increase in steatosis, hepatocellular ballooning or lobular inflammation using NASH CRN scoring system, from Baseline to the end of the DB phase

7.6.3. Multiple Imputation Analysis

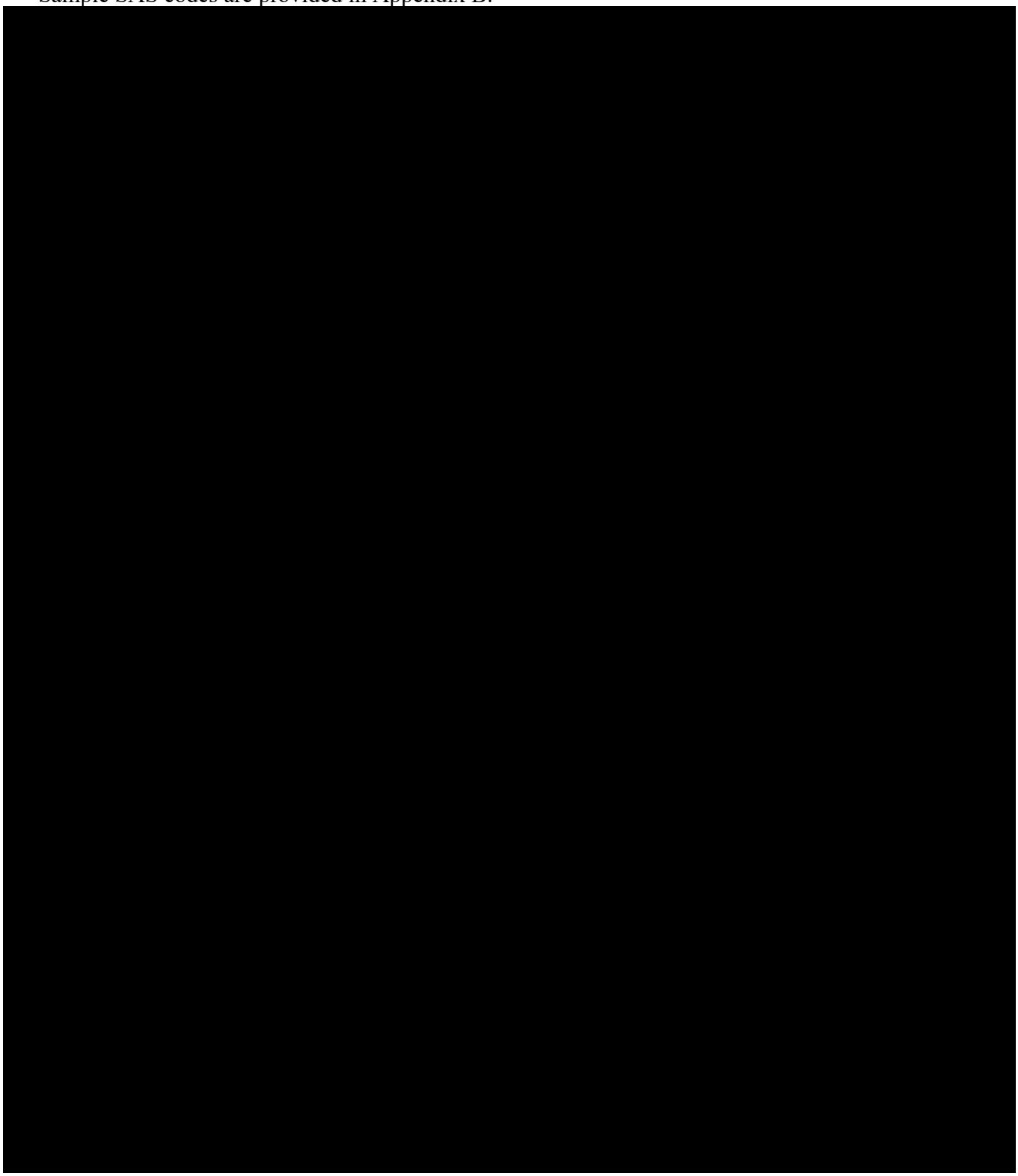
As discussed in Section 7.1, a multiple imputation method will be performed to estimate the expected treatment effect on the ITT population for the primary endpoint. To further assess the impact on these missing biopsy data, an additional sensitivity analysis will be conducted which will impute all missing primary endpoint using the same method, regardless of discontinuation reasons.

MAR assumption states that the missing data mechanism is independent of the missing observations conditional on the observed data. If missing outcome values are deemed to be MAR, and to avoid loss of efficiency, missing outcome values will be imputed using below multiple imputation steps.

1000 imputed data sets will be drawn separately for each randomized group, replacing missing outcome values with simulated values using predictive mean matching from a set of imputation models containing all potential prognostic baseline covariates as well as efficacy endpoints.

The weighted average method will be fitted to each of the 1000 imputed data sets and the results will be combined using Rubin's rules (Rubin, 1987).

Sample SAS codes are provided in Appendix B.



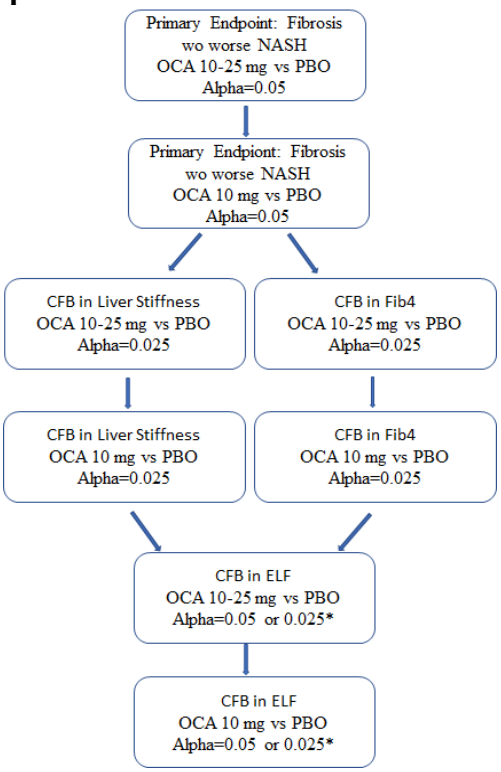
[redacted]

7.6.5. Multiple Comparisons/Multiplicity

The multiple testing procedure (MTP) based on gatekeeping procedure to control the family-wise type I error rate at 5% will be used in this study. The primary efficacy endpoint will be tested first at an alpha of 5% for the comparison between OCA 10 mg → 25 mg titration arm versus Placebo. If the primary hypothesis between OCA 10 mg → 25 mg titration versus placebo is rejected, the comparison between OCA 10 mg versus Placebo will be tested at an alpha of 0.05; otherwise, testing of the primary efficacy endpoint at the 10 mg dose will be considered descriptive and nominal. If the primary null hypothesis for both treatment arms are rejected, the gate keeping procedure will continue and test the hypotheses on the two secondary endpoints, respectively, at the same level, ie, change from baseline in liver stiffness, and change from baseline in Fib-4 will be tested at the same level between OCA 10 mg → 25 mg titration versus placebo.

The following chart (Figure 2) describes the step by step:

Figure 2: Multiple Comparisons Chart



CFB = change from baseline, ELF = enhanced liver fibrosis.
* Alpha level will be either 0.05 or 0.025, depending on the testing results from higher rank.

Multiple testing will follow a hierarchical order: the lower level will be tested only if the upper level is statistically significant at a pre-specified alpha level. OCA 10 mg vs PBO will be tested only if corresponding comparison for OCA 10-25 mg vs PBO is statistically significant. Nominal p-values will be reported for other secondary endpoints.

7.6.6. Efficacy Subgroups

Subgroup analyses will be performed only if there are enough subjects in each group (eg, >5 subjects per category).

1. Subgroup analyses for the primary endpoint using ITT and limited to summary statistics with no reporting of p-values will be presented for the following factors:
   a. Baseline diabetes status at enrollment: Yes, No
   b. BMI (<30 kg/m^2, ≥30 kg/m^2 and <35 kg/m^2, ≥35 kg/m^2)
   c. Age categories: <65, ≥65 years, ≥75 years
   d. Gender: Male, Female
   e. Race: White, Non-White (Black or African American, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native)
   f. Geographic Region: North America, Europe, and the Rest of the World (ROW)
   g. [redacted]
   h.
   i.
   j.
   k.
2. The change from baseline in the following parameters will be analyzed by response to the primary endpoint and treatment group. These summaries will be based on subjects with a post-baseline biopsy.
   [redacted]
   b. Lipids: LDL, HDL, VLDL and total cholesterol
   [redacted]
   - FIB-4, [redacted], ELF, and [redacted]

   [redacted]
   e. PD biomarkers of FXR activation including: C4, FGF-19, plasma bile acids
3. In the OCA 10 mg → 25 mg titration arm, there are subjects who either do not meet criteria for uptitration OR do not undergo uptitration for safety reason based on PI discretion. Two subgroup analyses will be performed:
   a. Comparison of Placebo and OCA 10 mg treatment group by adding those subjects who remained on OCA 10 mg but were randomized to the OCA 10 mg → 25 mg group.
   b. Comparison of Placebo and OCA 10 mg → 25 mg treatment group by excluding those subjects who remained at 10 mg in the OCA 10 mg → 25 mg group.

7.6.7. COVID-19

Due to the limitations caused by the COVID-19 pandemic, subjects who have been unable to complete the liver biopsy procedure at the Month 18 Visit, investigational product may be extended beyond the Month 18 Visit for a maximum of 3 additional months of the DB phase. All post-baseline biopsies obtained by the end of DB phase will be included in ITT population.

Due to the COVID-19, subjects who cannot have lab assessments performed at a central laboratory may have lab assessments performed at a local laboratory. The local lab results may be used if central lab assessment is not available.

To evaluate whether COVID-19 impacts on the study safety assessments, frequencies of selected adverse events summaries for the safety population 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. A summary of AEs will be provided.

8. EFFICACY ANALYSES IN OLE

A subset of efficacy assessments conducted in the DB phase will be repeated in the OLE Phase. Descriptive summaries will be provided by treatment group and scheduled visit.

Treatment groups will be defined as follows:

- Placebo / OCA 10 mg
- Placebo / OCA 10 mg → 25 mg titration
- OCA 10 mg / 10 mg
- OCA 10 mg → 25 mg titration / 10 mg → 25 mg titration

9. SAFETY ANALYSES IN DB AND OLE

DB Phase:

Safety summaries will be conducted using the safety population by treatment groups, subgroups, and total OCA (both OCA 10 mg and OCA 10 mg → 25 mg combined) during the DB phase.

If a subject is randomized to placebo and subsequently receives OCA, each safety assessment will be attributed to the administered OCA dose.

Treatment groups and subgroups will be defined as follows:

- Placebo
- OCA 10 mg
- OCA 10 mg → 25 mg: Subjects receiving this treatment will be further categorized into the following treatment subgroups:
  - Remained on 10 mg: These comprise subjects who do not titrate to the OCA 25 mg dose level, either because of ineligibility per protocol or any other reason. Such subjects should have received only OCA 10 mg for the entire duration of their participation in the DB phase.

  - Titrated to 25 mg: These comprise subjects who receive OCA 25 mg at any time after the titration eligibility assessment is made, even if they subsequently down titrated for any reason.
  - OCA 10-25 mg Combined
- Total OCA: both OCA 10 mg and OCA 10 mg → 25 mg combined

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

Selected safety summaries will be prepared including all subjects in the ITT Efficacy Population. Analyses based on the DB baseline will be performed using the actual treatment received in the DB phase.

OLE:

Adverse events and concomitant medications will be summarized cumulatively in OLE. i.e.

For exposure-adjusted incidence of adverse events, the summary will be based on the duration of active treatment.

For quantitative parameters including, but not limited to, vital signs, ECG, and clinical laboratory results, changes from the DB baseline will be presented by OLE nominal visits. Change from OLE baseline value may be additionally considered if deemed appropriate.

Treatment groups and subgroups will be defined as follows:

- Placebo / OCA 10 mg
- Placebo / OCA 10 mg → 25 mg titration: Subjects receiving this treatment will be further categorized into the following treatment subgroups:
  - Remained on 10 mg: These comprise subjects who do not titrate to the 25 mg dose level, either because of ineligibility per protocol or any other reason. Such subjects should have received only OCA 10 mg for the entire duration of their participation in the OLE phase.
  - Titrated to 25 mg: These comprise subjects who receive 25 mg at any time after the titration eligibility assessment is made, even if they subsequently down titrated for any reason.
  - OCA 10-25 mg Combined
- OCA 10 mg
- OCA 10 mg → 25 mg titration: Subjects receiving the same treatment as they received in DB phase:
  - Remained on 10 mg during DB phase: Subjects should have received only OCA 10 mg for the entire duration of their participation in the OLE phase.
  - Titrated to 25 mg during DB phase: Subjects should have received only OCA 25 mg for the entire duration of their participation in the OLE phase.
  - OCA 10-25 mg Combined during DB phase

9.1. Extent of Exposure

The exposure analysis will be presented separated for DB and OLE.

For DB phase, the extent of exposure will be summarized both continuously using descriptive statistics and categorically (n and % for ≤1 month, >1 and ≤3 months, >3 and ≤6 months, >6 and ≤9 months, >9 and ≤12 months, >12 and ≤18 months).

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, 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, 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 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:

- Compliance = duration of Exposure/(duration of Exposure + missed doses) *100

where drug holidays are not included in this calculation since compliance reflects subjects' adherence to the taking IP as required.

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.

9.2. Adverse Events

AE data will be collected from the time that signed informed consent is obtained until the subject fully completes her/his participation in the study. AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA, version 23.0). AEs are graded for severity (ie, intensity) using the Common Terminology Criteria for Adverse Events (CTCAE, Version 4.03).

All AE summaries will be restricted to TEAEs. TEAEs are defined as any AEs that newly appear, increase in frequency, or worsen in severity after 1 dose of investigational product during

the DB phase. 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 23.0. Each AE summary will be displayed by treatment group. Summaries that are displayed by SOC and PT will be ordered defined by descending order of incidence (OCA 10 mg → 25 mg titration, followed by OCA 10 mg, followed by placebo, detail in Section 9) of SOC and PT 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 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 SOC and PT.
- Subject incidence of TEAEs leading to investigational product withdrawal by SOC and PT.
- Subject incidence of TEAEs by SOC, PT, and highest severity (common terminology criteria for adverse events [CTCAE] grade).

  Notes: 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 SOC and PT.

  Notes: 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 SOC and PT.

  Notes: 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 SOC and PT.

  Notes: 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 SOC and PT with frequency at least 5% within any treatment group.
- Subject incidence of TEAEs by SOC, PT, and closest relationship to liver biopsy (Related/Not Related).

Notes: 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 SOC and PT.
- Overall Summary of TEAEs Over Time
- Exposure-adjusted incidence of TEAEs leading to death
- Exposure-adjusted incidence of serious TEAEs
- Exposure-adjusted incidence of TEAEs leading to investigational product withdrawal

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

- All TEAEs
- Serious TEAEs

  Notes: This is a subset of the TEAEs where serious is marked as "Yes".
- CTCAE Grade 3 or higher TEAEs

  Notes: This is a subset of TEAEs where severity is marked as CTCAE grade 3, 4, or 5)
- Study Drug Related TEAEs

  Notes: This is a subset of the TEAEs where relationship to study drug marked as "Definite," "Probable," or "Possible".
- Liver Biopsy-related TEAEs

  Notes: This is a subset of the TEAEs where relationship to liver biopsy marked as "Definite", "Probable", or "Possible".
- TEAEs leading to Study Drug Withdrawal
- TEAEs leading to study discontinuation
- TEAEs leading to Death

  Notes: This is a subset of the TEAEs where outcome is indicated as "Fatal" or the CTCAE grade is 5.

9.3. Adverse Events of Special Interest

Adverse Events of Special Interest (AESI) are defined using SMQs and/or custom MedDRA version 23.0 for 747-304. The following adverse events of special interest (AESIs) will be evaluated and summarized for Safety populations.

- Pruritus
- Hepatic adverse events
- Cardiovascular adverse events (CV is addressed in detail in Section 9.4)
- Dyslipidemia

- Gallbladder disease and related complications
- Pancreatitis
- Renal injury
- Urolithiases
- Hyperglycemia /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 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 treatment-emergent AESIs by PT
- Exposure-adjusted incidence of serious treatment-emergent AESIs by PT
- Exposure-adjusted incidence of treatment-emergent AESIs by baseline diabetes status
- Exposure-adjusted incidence of severe (Grade 3+) AESIs by PT
- KM statistics for time to first onset of treatment-emergent AESI
- KM statistics for time to onset of first severe (Grade 3+) or serious 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 each AESI, analyses will be performed separately for the AESI collected on-treatment only and during study. Exposure-adjusted incidence rate difference (with corresponding 95% CIs) will be calculated for the number of subjects with any event by subcategory of interest (not for each PT).

Additional special analyses for each AESI may be provided if deemed necessary.

9.3.1. 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 Population.

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 last known date during the study minus the first dose date plus 1 divided by 365.25 days/year.

In addition, an analysis of treatment-emergent pruritus events by preferred term and severity over time will be performed.

- 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) used to treat the pruritus, adjustments to investigational product dosing frequency or amount, mandatory investigational product interruption, and/or holiday. The time between each event and it's preceding one will be presented per subject.

9.3.2. Hepatic Adverse Events

Hepatic adverse events 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
- pregnancy-related hepatic disorders

9.3.3. Cardiovascular Adverse Events

Cardiovascular adverse events are defined as:

- Embolic and thrombotic events broad SMQ
- Ischemic heart disease broad SMQ
- Central nervous system vascular disorders narrow SMQ

9.3.4. Dyslipidemia

Dyslipidemia is defined as events included in the dyslipidemia SMQ.

9.3.5. Gallbladder Disease and Related Complications

Gallbladder disease and related complications adverse events are defined as:

- Gallbladder related disorders narrow SMQ
- Gallstone related disorders narrow SMQ
- Additional Preferred terms 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, or Perforation bile duct

9.3.6. Pancreatitis

Pancreatitis is defined as acute pancreatitis narrow SMQ.

9.3.7. Renal Injury

Renal injury is broadly defined as:

- Acute Renal Failure SMQ
- Chronic Kidney Disease SMQ
- Proteinuria SMQ
- Renovascular disorders broad SMQ
- Tubulointerstitial disease broad SMQ

9.3.8. Urolithiases

Urolithiases are defined as high-level group term Urolithiases (under the SOC Renal and Urinary Disorders).

9.3.9. Hyperglycemia/New Onset Diabetes Mellitus

Hyperglycemia/new onset diabetes mellitus are defined as hyperglycemia/new onset diabetes mellitus narrow SMQ.

9.4. Cardiovascular Risk

Summaries of adjudicated cardiovascular events include core MACE (cardiovascular death, non-fatal MI, non-fatal stroke) and expanded MACE will be performed as follows, in addition to the AESI analysis in Section 9.3.3.

9.4.1. Adjudicated Cardiovascular Events

Summaries of adjudicated cardiovascular events including major adverse cardiovascular events (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.

9.4.2. Markers of Cardiovascular Safety

Framingham Risk Score (FRS) will be summarized by treatment group using descriptive statistics of the results at baseline and at each post-baseline visit. The change and percentage change from baseline values will be summarized. All summaries will be based on observed values. A shift table from baseline FRS to each post-baseline visit will also be provided.

9.4.3. Serum Chemistry Lipids

Lipoproteins (LDL, HDL, VLDL, and triglycerides, ApoB, ApoA-1, ApoE, Lp[a]), total cholesterol, triglycerides, and PCSK9 results by treatment group using descriptive statistics at baseline and at each on-study evaluation presented. Laboratory values, change, and percentage change from baseline will also be summarized.

9.5. Adjudicated Acute Kidney Injury (AKI) Events

Summaries of adjudicated acute kidney injury events will include the overall summary of AKI TEAEs, incidence of AKI TEAEs, incidence of serious AKI TEAEs, annualized adjudicated AKI event rate, and a time-to-event analysis. All summaries of incidence will include the

associated exact binomial 95% confidence interval (CI). Summaries will be carried out for the Safety population.

9.6. Adjudicated Hepatic Injury (DILI) Events

Summaries of adjudicated hepatic injury events will include the overall summary of hepatic injury TEAEs, incidence of hepatic injury TEAEs, incidence of serious hepatic injury TEAEs, annualized adjudicated hepatic injury event rate, and a time-to-event analysis. All summaries of incidence will include the associated exact binomial 95% confidence interval (CI). Summaries will be carried out for the Safety population.

Potential DILI Triggers: Potential cases of drug induced liver injury (DILI) 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). As part of a comprehensive DILI analysis, the HSAC will apply these criteria to all subjects in 747-304.

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

9.7. Clinical Laboratory Evaluations

Laboratory parameters will be summarized in both conventional units and the standard international (SI) system of units. The submitted datasets will also include results from local lab evaluations.

All clinical laboratory data will be presented in by-subject data listings. The following analyses will be done for the Safety population.

9.7.1. 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 will be presented. 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.

9.7.2. Hepatic Biochemical Analysis

[redacted]

Summary statistics will be presented by visit, including change from baseline and percent change from baseline.

While portal hypertension is defined by hepatic venous pressure gradient, there are non-invasive surrogate markers that can rule-out clinically significant portal hypertension (CSPH) as well as identify the potential for increased risk of portal hypertension.

- A shift summary for platelet (PLT) count by baseline category: 100,000 - 150,000/mm3, 150,000 -200,000/mm3 and >200,000/mm3 that are improved, stable, or worsened (worse post-baseline value) at month 6, 12, and 18 visits will be presented.
- In addition, a composite result of baseline FibroScan (FS) and PLT count (to rule-out the presence of CSPH), will be defined as follows:
  - Composite Result = "Yes" if FS<=15kPa and PLT>=150,000/mm^3 – No evidence of CSPH
  - Otherwise, composite result = "No"
  - A shift summary for a composite result improved (N->Y), stable (Y->Y and N->N), worsened (Y->N) from baseline to month 6, 12, and 18 visits will also be presented.

Evaluation of drug-induced serious hepatotoxicity (eDISH) will be conducted to visually compare maximum [redacted] values at any post-baseline visit. For subjects with baseline [redacted] 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 [redacted] Plots based on multiples of elevations above the baseline value will use [redacted]

The summary table as well as visual plots will be based on central lab data.

In addition, eDISH using maximum post-baseline [redacted] will be repeated with central and local lab data.

9.7.3. 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.

Only descriptive statistics will be provided.

9.7.4. Urine Chemistry and Urinalysis

Quantitative urine chemistry and urinalysis results by treatment group and by baseline body weight will be summarized 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.

9.7.5. 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 and serum creatinine. The summary will also be repeated for baseline CKD stage.

Annualized change in eGFR (mL/min/1.73 m^2/yr) will be summarized descriptively. The proportion of subjects that experience rapid eGFR decline (defined as >4 mL/min/1.73 m^2/yr) will be presented.

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

9.8. Pregnancies

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

9.9. Vital Signs

The results, change from baseline, and percentage change from baseline to each on-study evaluation visit will be summarized for systolic blood pressure, diastolic blood pressure, sitting heart rate, respiratory rate, body weight, waist circumference and BMI.

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

9.10. Electrocardiograms

Summary statistics for the values, change from baseline, and percent change from baseline to each on-study evaluation visit will be presented for RR interval, PR interval, QRS duration, QT interval, QTcF interval. A summary of abnormal ECG results that are identified by centrally analyzed ECG results by ERT will be provided.

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.

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

9.11. Safety Subgroups

Subgroup analyses for safety selected endpoints including overall TEAE, serious TEAEs, overall treatment-emerging AESI, and serious treatment-emerging AESI may be presented for the Safety Population and may include the following factors:

a. Baseline diabetes status at enrollment: Yes, No
b. Age categories: <65, ≥65 years, ≥75 years
c. Baseline in BMI (<30 kg/m^2, ≥30 kg/m^2 and <35 kg/m^2, ≥35 kg/m^2)
d. Baseline Framingham (FRS) risk score: ≤10, >10 - ≤20, >20
e. Baseline eGFR

10. PK/PD ANALYSES IN DB PHASE

Details of PK analyses will be specified in a separate Clinical Pharmacology Analysis Plan (CPAP).

11. CHANGE FROM PROTOCOL PLANNED ANALYSIS

The following changes were made after the protocol 7 version was finalized:

1. Protocol Stat section 17.4 (page 139), secondary efficacy analyses will be conducted using ITT populations.
2. Impact: the secondary endpoints are in multiple testing and will be based in ITT population. Other mITT or PP population may be analyzed but they are sensitivity analysis.
3. Inclusion of other exploratory histological endpoints as measured by Ishak scoring system and AI-based histologic Measurement of NASH (AIM NASH).
4. The adverse events of special interest defined in the protocol are modified to align with the other NASH studies.

Original text

Adverse events of special interest include cardiovascular, pruritus, renal, urinary tracts stones including nephrolithiasis, gallbladder/gallstones-related pancreatitis, hepatic, dyslipidemia and hyperglycemia/new-onset diabetes mellitus AEs.

Change made to

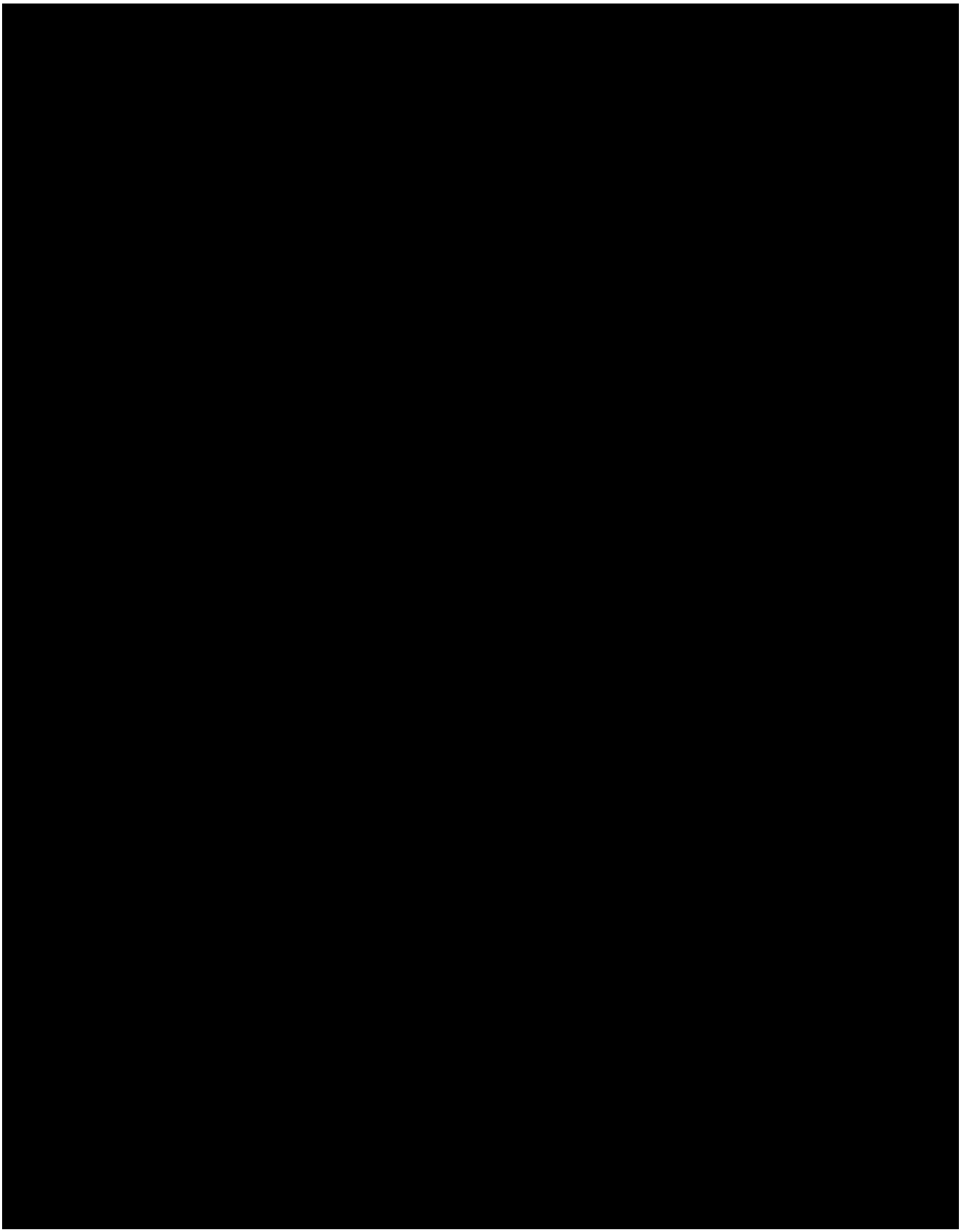
Adverse events of special interest include pruritus, hepatic adverse events, cardiovascular adverse events, dyslipidemia, gallbladder disease and related complications, pancreatitis, renal injury, urolithiases, hyperglycemia/new onset diabetes mellitus.

Additional changes made were in response to FDA's information request from 18 Jan 2022. Other notable change is the inclusion of liver non-invasive procedures to the secondary endpoints. Summary of key updates can be found in Section 1.

APPENDIX A. CHILD-PUGH SCORING SYSTEM

Factor	Units	Points		
		1	2	3
Serum total bilirubin	µmol/L	<34	34-50	>50
	mg/dL	<2.0	2.0-3.0	>3.0
Serum albumin	g/L	>35	28-35	<28
	g/dL	>3.5	2.8-3.5	<2.8
Prothrombin time	Seconds prolonged	0-3	4-6	>6
	INR	<1.7	1.7-2.3	>2.3
Ascites		None	Mild	Moderate-Severe
Hepatic encephalopathya		None	Grade 1 or 2	Grade 3 or 4

INR = international normalized ratio
Child-Pugh criteria: Pugh 1973, Lucey 1997.
West Haven criteria: Vilstrup 2014.
[a] Grade 0: normal consciousness, personality, neurological examination, electroencephalogram
Grade 1: restless, sleep disturbed, irritable/agitated, tremor, impaired handwriting, 5 cps waves
Grade 2: lethargic, time-disoriented, inappropriate, asterixis, ataxia, slow triphasic waves
Grade 3: somnolent, stuporous, place-disoriented, hyperactive reflexes, rigidity, slower waves
Grade 4: unrousable coma, no personality/behavior, decerebrate, slow 2-3 cps delta activity



APPENDIX C. LIST OF LAB TESTS USE MEAN OF PRE-DOSE VALUES AS BASELINE

ActiTest Score	Ery. Mean Corpuscular Volume
Activated Partial Thromboplastin Time	Erythrocytes
Alanine Aminotransferase	Ferritin
Albumin	Free Fatty Acid
Albumin	Haptoglobin
Albumin/Creatinine	Hematocrit
Alkaline Phosphatase	Hemoglobin
Alkaline Phosphatase Serum or Plasma	Hemoglobin
Alk-Phosphatase Other Calc (ARUP)	Ketones
Alpha-2 Macroglobulin	Leukocyte Esterase
Apolipoprotein A1	Leukocytes
Apolipoprotein B	Liver Specific Alkaline Phosphatase
Apolipoprotein E	Lymphocytes
AshTest Score	Lymphocytes/Leukocytes
Aspartate Aminotransferase	Magnesium
Basophils	Monocytes
Basophils/Leukocytes	Monocytes/Leukocytes
Bicarbonate	Neutrophils
Bilirubin	Neutrophils/Leukocytes
Bilirubin	Nitrite
Blood Urea Nitrogen	pH
Bone Specific Alkaline Phosphatase	Phosphate
C Reactive Protein	Platelets
Calcium	Potassium
Chloride	Protein
CK-18 (M30)	Protein
CK-18 (M65)	Prothrombin Intl. Normalized Ratio
Creatine Kinase	Prothrombin Time
Creatinine	PT Ratio
Creatinine	Sodium
Direct Bilirubin	Specific Gravity
Eosinophils	Thyrotropin
Eosinophils/Leukocytes	Thyroxine, Free
Ery. Mean Corpuscular Hemoglobin	TIMP-1 for ELF Score
Ery. Mean Corpuscular HGB Concentration	Triiodothyronine, Free