Statistical Analysis Plan for Protocol 747-308

A Randomized, Double-blind, Placebo-controlled, Phase 2/3 study to Assess the Efficacy, Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of Obeticholic Acid Compared to Placebo in Pediatric Subjects with Biliary Atresia, Post-Hepatoportoenterostomy

OBETICHOLIC ACID (OCA)

Protocol Version and Date:	Version 1.0 – 8/31/23 Version 1.0 (China 01) – 12/19/23 Version 1.0 (EU 01) – 3/22/24 Version 1.0 (UK 01) – 5/16/24 Version 1.0 (Vietnam 01) – 12/19/23
Phase:	Phase 2/3
Author:	[redacted]
Sponsor:	Intercept Pharmaceuticals, Inc. 305 Madison Avenue Morristown, NJ 07960 Tel: 646-747-1000 Fax: 858-558-5961
Sponsor Representative	
Analysis Plan Date:	21 OCT 2025
Analysis Plan Version:	1.0

Approvals:

Authored by:

[redacted]
Electronically signed by: [redacted]
Reason: I am the author of this document
Date: Oct 21, 2025 19:27:15 GMT+2

21-Oct-2025
Date

[redacted] Global Biometric Sciences – Biostatistics
Alfasigma S.p.A.

Approved by:

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 21, 2025 11:31:28 PDT

21-Oct-2025
Date

[redacted], Clinical Research
Alfasigma S.p.A.

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 21, 2025 15:00:48 EDT

21-Oct-2025
Date

[redacted] (on behalf of the head of Global Biometric Sciences)
[redacted] Global Biometric Sciences
Alfasigma S.p.A.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS 6
1. VERSION HISTORY 9
2. INTRODUCTION 10
3. STUDY OBJECTIVES AND ESTIMANDS 10
3.1. Study Objectives 10
3.1.1. Primary Objective 10
3.1.2. Secondary Objectives 10
3.1.3. Exploratory Objectives 11
3.2. Estimands 11
3.2.1. Primary Estimand(s) 11
3.2.2. Primary Estimand(s) for Secondary Endpoints 12
3.2.3. Additional Estimand(s) 12
4. SUMMARY OF STUDY DESIGN 12
4.1. Randomization 14
4.2. Blinding and Unblinding 14
4.3. Sample Size Determination 15
5. ANALYSIS POPULATIONS AND APPROACHES TO ANALYSIS 16
6. GENERAL STATISTICAL METHODS 16
6.1. Hypothesis and Decision rules 16
6.1.1. Primary Hypothesis 16
6.1.2. Secondary Hypotheses 16
6.2. General Methods 17
6.2.1. Analysis for Binary Endpoints 17
6.2.2. Analysis for Continuous Endpoints 17
6.2.3. Analysis for Time-to-Event Endpoints 17
6.3. Methods to Manage Missing Data 18
6.3.1. Analysis for Time-to-Event Endpoints 18
6.3.2. Analysis for Binary Endpoints 18
6.3.3. Analysis for Continuous Endpoints 18
6.3.4. Analysis for Categorical Endpoints 18
7. SUBJECT DISPOSITION 18

8. PROTOCOL DEVIATIONS 19
9. DEMOGRAPHICS AND BASELINE DISEASE CHARACTERISTICS 19
10. PRIOR THERAPY AND MEDICAL HISTORY 21
11. EFFICACY ANALYSES 21
11.1. Primary [Efficacy] Endpoint 21
11.2. Analyses for Time-to-Event Endpoints 21
11.2.1. Primary Analysis 21
11.2.2. Sensitivity/Supplementary Analyses 22
11.3. Secondary [Efficacy] Endpoints 22
11.3.1. Composite Endpoint of Biomarkers 22
11.3.2. Biomarkers of Hepatobiliary Function 23
11.3.3. Plasma Fat Soluble Vitamins D and K 23
11.3.4. Transient Elastography (TE) 23
11.4. Exploratory [Efficacy] Endpoints 23
11.5. Baseline Variables 23
11.6. Multiple Comparisons/Multiplicity 24
11.7. Subgroup Analyses of Efficacy 24
12. SAFETY ANALYSES 24
12.1. Exposure to Study Treatment 24
12.2. Adverse Events 25
12.2.1. Treatment-emergent Adverse Events 26
12.2.2. Serious Adverse Events 26
12.2.3. Adverse Events of Interest 27
12.2.4. Treatment-emergent Adverse Events Analyses 27
12.2.5. Missing or Partial Date Handling 28
12.3. Clinical Laboratory Evaluations 29
12.3.1. Hematology 30
12.3.2. Serum Chemistry 30
12.3.3. Other Laboratory Parameters 30
12.4. Anthropomorphic Measurements and Vital Signs 30
12.5. Electrocardiograms (ECGs) 31
12.6. Prior and Concomitant Medications 31

12.7. Subgroup Analyses of Safety....31
13. INTERIM ANALYSES....31
13.1. Introduction....31
13.2. Interim Analyses and Summaries....32
14. PHARMACOKINETIC ANALYSES....32
15. PHARMACODYNAMIC ANALYSES....33
16. REFERENCES....34
APPENDIX A. DATA PRESENTATION CONVENTIONS:....35
APPENDIX B. DEFINITION AND USE OF VISIT WINDOW IN REPORTING....36
APPENDIX C. STATISTICAL METHODOLOGY DETAILS....39

LIST OF TABLES

Table 1: Summary of Changes....9
Table 2: Primary Estimand....11
Table 3: Primary Estimand for Composite Biomarker Endpoint....12
Table 4: Primary Estimand for Liver Stiffness....12
Table 5: Supplemental Estimands to the Primary Estimand....22
Table 6: Bile Acid Totals and Individual Bile Acids with Molecular Weights....33
Table 7: Bile Acid Totals and Individual Bile Acids with Molecular Weights....33
Table 8: Analysis Visits....37

LIST OF FIGURES

Figure 1: Study Schematic....13

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or Specialist Term	Explanation
AE	Adverse Event
AED	Adult Equivalent Dose
AEI	Adverse events of interest
ALP	Alanine phosphatase
ALT	Alanine amintotransferase
ANCOVA	Analysis of Covariance
APRI	Aspartate aminotransferase to platelet ratio
AR(1)	Autoregression model 1
AST	Aspartate aminotransferase
ATC	Anatomic Therapeutic Chemical
BMI	Body Mass Index
BUN	Blood urea nitrogen
CAR	Censoring at Random
CI	Confidence interval
CMH	Cochran Mantel Haenszel
CS	Clinically significant
CRO	Clinical Research Organization
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DB	Double-Blind
DOB	Date of birth
DOIC	Date of informed consent
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
EOS	End of study
FXR	Farnesoid X receptor
GGT	Gamma-glutamyl transferase
HDL	High density lipoprotein
HPE	Hepatoportoenterostomy
IC	Intercurrent events

Abbreviation or Specialist Term	Explanation
ICH	International Council for Harmonization
IEC	Independent Ethics Committee
INR	International normalized ratio
IQR	Interquartile range
IRB	Institutional Review Board
ITT	Intent-To-Treat
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LDL	Low densitiy lipoprotein
LOCF	Last observation carried forward
LS	Least squares
MCH	Mean corpuscular hemoglobin
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean cell volume
MedDRA	Medical Dictionary of Regulatory Activities
MELD-NA	Model of End-stage Liver Disease with sodium
MMRM	Mixed model for repeated measures
MI	Multiple imputation
mITT	Modified Intention-to-Treat
NCS	Not clinically significant
NEC	Not elsewhere classified
NRI	Non-Responder Imputation
OC	Observed case
OCA	Obeticholic Acid
PBC	Primary biliary cholangitis
PBO	Placebo
PD	Pharmacodynamics
PELD	Pediatric End-stage Liver Disease
PIP	Pediatric Investigational Plan
PK	Pharmacokinetics
PT	Preferred term
PTT	Prothrombin time

Abbreviation or Specialist Term	Explanation
QTcF	QT interval corrected by Fridericia's formula
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard Deviation
SE	Standard Error of the mean
SOC	System organ class
SMQ	Standardized MedDRA queries
SUSAR	Suspected unexpected serious adverse reaction
TSH	Thyroid stimulating hormone
TEAE	Treatment-emergent adverse event
UDCA	Ursodeoxycholic acid
ULN	Upper limit of normal
VAS	Visual analogue scale
VLDL	Very low density protein
WHO	World Health Organization

1. VERSION HISTORY

Table 1: Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
1.0 15 SEP 2025	Original 31 Aug 2023	N/A	N/A

2. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analyses and data presentations planned for protocol 747-308 version 1.0 dated 31 Aug 2023, Version 1.0 (China 01) 12/19/23, Version 1.0 (EU 01) 3/22/24, Version 1.0 (UK 01) 5/16/24, Version 1.0 (Vietnam 01) 12/19/23. This study is a Randomized, Double-Blind, Placebo controlled, Phase 2/3 Study to Assess the Efficacy, Safety, Tolerability, Pharmacokinetics and Pharmacodynamics of Obeticholic Acid Compared to Placebo in Pediatric Subjects with Biliary Atresia, Post-hepatoportoenterostomy. This SAP provides a detailed description of the strategies, rationales, and statistical techniques to be used to meet the study objectives and additional details to the statistical analyses that were described in the protocol. This SAP will be finalized and signed prior to any formal interim analysis (IA) or database lock, whichever comes first. 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* in the CSR.

3. STUDY OBJECTIVES AND ESTIMANDS

3.1. Study Objectives

3.1.1. Primary Objective

To evaluate the effect of OCA compared to placebo in conjunction with established local standard of care on clinical outcomes in subjects with biliary atresia who have had a successful Kasai procedure as measured by time to first occurrence of any of the following adjudicated events, derived as composite event endpoints:

- Death (all-cause)
- Liver transplant
- Pediatric end-stage liver disease (PELD) ≥17/MELD-NA ≥15
- Hospitalization (as defined by a stay of 24 hours or greater) for new onset or recurrence of:
  - Variceal bleed
  - Hepatic encephalopathy (as defined by a West Haven score of ≥2)
  - Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis)
- Clinically evident ascites related to portal hypertension (diuretic-resistant ascites requiring therapeutic paracentesis at a frequency of at least twice in a month)

3.1.2. Secondary Objectives

To evaluate the:

- PK of OCA as measured by unconjugated plasma OCA (parent), glyco-OCA, tauro-OCA, and total OCA (molar sum of OCA and its active conjugates).

- Pharmacodynamics (PD) of OCA as measured by biomarkers of FXR activation: FGF-19, C4, and endogenous bile acids, and
- Biomarkers of hepatobiliary function: GGT, total and direct (conjugated) bilirubin
- Effect of OCA on liver stiffness by transient elastography (if available at site)
- Disease progression as measured by plasma levels of fat-soluble vitamins (Vitamins D and K)
- Safety and tolerability of OCA

3.1.3. Exploratory Objectives

To evaluate the

- Acceptability and swallowability of the dosing regimen
- Palatability of the dosing regimen

3.2. Estimands

3.2.1. Primary Estimand(s)

Table 2: Primary Estimand

Estimand		
Population	Pediatric patients with biliary atresia post-HPE Intent-to-Treat (ITT)	
Variable/Endpoint	Time from randomization to the first occurrence of any of the following adjudicated post-randomization events, derived as composite endpoint: 1. Death (all cause) 2. Liver transplant 3. Pediatric end-stage liver disease (PELD) score >= 17/model of end-stage liver disease (MELD-NA) >= 15 4. Hospitalization (>= 24 hours) for new onset or recurrence of 1. Variceal bleed 2. Hepatic encephalopathy (West Haven score >= 2) 3. Spontaneous bacterial peritonitis (confirmed by diuretic-resistant paracentesis) 5. Clinically relevant ascites related to portal hypertension (ascites requiring therapeutic paracentesis at a frequency of at least twice a month)	
Population level summary	Log rank test of clinical outcomes adjusted for age (using median as cut-off) of the OCA group versus placebo group on the time to first occurrence of the clinical endpoint (see above). Subjects who do not experience an event will be censored at the time of their last contact.	
Intercurrent events (IC) and strategy	Prohibited medication and treatment discontinuation	Treatment policy: treatment effect regardless of the IC

3.2.2. Primary Estimand(s) for Secondary Endpoints

Table 3: Primary Estimand for Composite Biomarker Endpoint

Estimand		
Population	Pediatric patients with biliary atresia post-HPE Intent-to-Treat (ITT)	
Variable/Endpoint	Biomarker responder if any of the below conditions is met: 1. Percent change from baseline GGT <= -40% 2. Percent change in bilirubin <= -25% At month 4 (first interim analysis) and final assessment (final analysis)	
Population level summary	Cochran Mantel Haenzel (CMH) test adjusted by age group (using median age as cut-off value). Subjects with missing values will be regarded as non-responders.	
Intercurrent events and strategy	Prohibited medication and treatment discontinuation	Treatment policy: treatment effect regardless of the IC

Table 4: Primary Estimand for Liver Stiffness

Estimand		
Population	Pediatric patients with biliary atresia post-HPE Intent-to-Treat (ITT)	
Variable/Endpoint	Median Liver Stiffness as measured by transient elastography at Week 58	
Population level summary	Wilcoxon rank sum test. The median difference between treatment groups will be constructed using the Hodges-Lehman estimate. If the differences between treatment groups are not approximately symmetrical around the median, the Sign Test may be used instead.	
Intercurrent events and strategy	Prohibited medication and treatment discontinuation	Treatment policy: treatment effect regardless of the IC

3.2.3. Additional Estimand(s)

See Section 11.2.2.

4. SUMMARY OF STUDY DESIGN

This is a phase 2/3 study in subjects with biliary atresia with successful HPE as defined in Section 8 of the protocol. The study is designed to support the Pediatric Investigational Plan (PIP).

Screening phase: Subjects will be screened during a 14-day screening period. Subjects must be able to swallow the placebo tables to be included in the study. Thus, subjects will receive placebo tablet(s) at the Screening Visit to evaluate acceptability and the ability to swallow. The study will initially enroll pediatric subjects aged 2 to <18 years to be randomized to receive OCA or placebo in a DB manner. Following the interim analysis and assessment of safety in older children, enrollment of subjects <2 years of age will be considered. Enrollment of subjects <2 years of age will only start after regulatory authority approval via a protocol amendment.

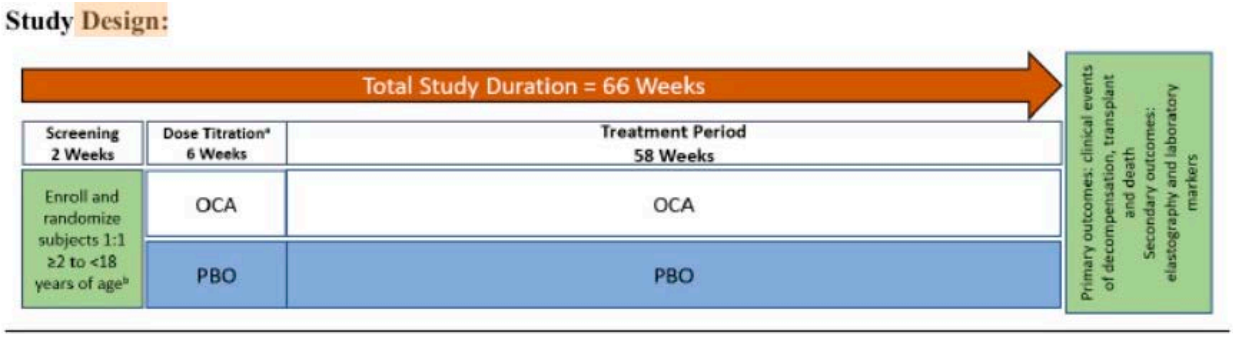
Double-Blind Treatment Phase

Dose Titration Phase: Subjects who satisfy the inclusion/exclusion criteria will be randomized to receive double-blinded placebo or OCA in a 1:1 ratio. Doses of study drug will be titrated every 2 weeks in a stepwise manner for the first 6 weeks, starting at 1.5 mg, through 3 mg AED to a maximum of 5 mg AED, as tolerated (section 7.2 of protocol). Subjects who do not tolerate higher OCA doses may return to their previous dose.

Age Expansion Treatment Phase: Following the 6-week titration, subjects will continue at the tolerated dose until Week 64, after which subjects will be followed-up for an additional 4 weeks.

The study schematic is given in Figure 1.

Figure 1: Study Schematic



AED=adult equivalent dose; OCA=obeticholic acid; PBO=placebo
Note: n=approximately 144 subjects
[a] Doses of investigational product (i.e., OCA or placebo) will be titrated every 2 weeks in a stepwise manner for the first 6 weeks, starting at 1.5 mg AED and titrating through 3 mg AED to a maximum of 5 mg AED, as tolerated.
[b] Subjects aged <2 years will be enrolled only if supported by safety and efficacy data from the first interim analysis.

Efficacy will be assessed using a composite endpoint of adjudicated liver-related events including all-cause death (see Section 3.1). Safety will be assessed using Adverse event reporting, laboratory safety, vital signs and ECG measurements.

A first interim analysis will be conducted when approximately 25 to 30 subjects reach their Month 4 visit. Assessments of GGT, total and direct bilirubin, total bile acid and bile acid totals, and levels of plasma OCA, OCA conjugates and total-OCA at available timepoints, will be evaluated. An unblinded Data Safety Monitoring Board (DSMB) will assess these results along with overall safety of the drug and will advise on 1) any safety concerns that would prohibit inclusion of subjects <2 years of age into the study, and 2) evidence of failure to sufficiently impact biomarkers, which would indicate futility of the study.

The timing of the second interim analysis will be approximately 12 months later (to be confirmed by the DSMB following the first interim analysis) to establish conditional power,

which may lead to stopping the study early. The DSMB will continue to oversee the data throughout the study.

4.1. Randomization

This study will be conducted in a DB, placebo-controlled manner. Enrolled subjects will be randomized in a 1:1 ratio to 1 of 2 treatment groups (OCA or matching placebo) based on a pre-defined randomization code (generated by designee) using an interactive web response system (IWRS). The IWRS will serve as the study drug 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 study drug (OCA or placebo) dispensing information will be provided (e.g., bottle numbers allocated) while maintaining the study blind.

4.2. Blinding and Unblinding

Study INT747-308 is a randomized, double-blind, and placebo-controlled study.

The subjects, Investigator, study site staff, the study team and adjudication committees will be blinded to the subject's treatment allocation during the subject's participation in the study.

Treatment assignment for individual subjects in the study will be made available to the Investigator for medical emergency use only (e.g., in situations where knowledge of the blinded treatment assignment is necessary for the treatment of an SAE) through the IWRS system. When possible, the Medical Monitor should be consulted if a medical emergency necessitates unblinding. If it is not reasonable to inform the Medical Monitor in advance of unblinding, the Investigator must promptly document the case and the rationale for unblinding in the subject's source record and should subsequently contact the Medical Monitor to discuss any premature unblinding of treatment assignment. Procedures for unblinding a subject's treatment will be provided separately to the Investigator.

Similarly, if the Medical Monitor breaks the blind for the purpose of evaluating a medical emergency or emergent safety issue, the Medical Monitor will document within the study source documentation the rationale, circumstances, and the person(s) being informed about the unblinding.

Access to randomization codes and corresponding treatment assignments will also be made available through the IWRS system to the designated Intercept staff responsible for unblinding SUSARs for reporting to the regulatory authorities.

In order to fulfil its responsibilities, the DSMB will be unblinded in its assessment of clinical trial data. The DSMB will have full access to all data as needed for safety assessment and will have access to comparative results of safety data, aggregated by treatment arm. The DMC can also request to be fully unblinded to individual patient treatment assignments at any point.

The DSMB will document details about any unblinded subject data reviews in closed session DSMB minutes, which will be made available to the Sponsor only after the double-blind database is locked, and the study is unblinded.

The Independent CRO will generate an Open Report that includes blinded information and a Closed Report that includes semi-blinded information (labeled A vs. B vs. C, for example).

Only the DMC and the Independent CRO will have access to the Closed Report.

The DSMB may direct questions and requests further data directly to the Independent CRO. If the independent CRO cannot complete these requests or requires clarification about the database structure, then further assistance from the Intercept team will be necessary. The DSMB will determine what is communicated to the Intercept team regarding additional requests. Any

communication between the Independent CRO and the Intercept team will be performed in a blinded fashion.

The unblinded data access to the DSMB interim analyses will be described in the Data Monitoring Committee charter for Intercept clinical trials.

Select study team members (i.e. data managers, biostatisticians, clinical pharmacologist) involved in regulatory or DSMB interactions may be unblinded to individual (per participant ID) and aggregate treatment allocation once participants are randomized. These select study team members will be distinct and independent from the team members involved directly with day-to-day study operations.

4.3. Sample Size Determination

The overall study sample size is estimated based on the clinical outcome composite endpoint analysis at the end of the study; time to the first occurrence of any component of the adjudicated clinical outcome (death, liver transplant, and markers of decompensated liver damage).

A total of 66 composite events will be required in the OCA and placebo groups in a 1:1 ratio, to provide about 80% power via two-sided log-rank test at the 0.05 significance level, assuming a hazard ratio of 0.50 between the OCA and placebo groups. Approximately 144 subjects for the 2 groups will be required to obtain the 66 events based on the additional assumptions:

- Total study duration of the study (first subject consented/randomized to last subject completing the last study visit of the DB phase) will be approximately 4 years, allowing 1.5 years accrual rate
- Placebo event rate of 52% over 1.5 years

This provides a sample of 105 subjects to obtain 66 events. Accounting for a 20% discontinuation rate, the sample size required to achieve the same number of events is 132. As this is a pediatric study and the discontinuation rate is unknown and to ensure enough events are observed, approximately 144 subjects are expected to be enrolled.

5. ANALYSIS POPULATIONS AND APPROACHES TO ANALYSIS

Data for all subjects will be assessed to determine if subjects meet the criteria for inclusion in each analysis population prior to unblinding and releasing the database and classifications will be documented per standard operating procedures.

Population	Description
Enrolled	All subjects who sign the informed consent Form.
Intent-to-treat (ITT)	The ITT Population will include all randomized subjects. The treatment assignment will be based on randomized treatment. The ITT Population will be the primary population used for all interim and final efficacy analyses and PD analyses, unless stated otherwise. Subjects will be analyzed according to the randomized intervention.
mITT or Evaluable	The mITT Population will include all subjects from the ITT Population, who took at least one dose of study drug (OCA or placebo) and who received treatment for at least 4 months. The mITT population will serve as the secondary analysis population for the efficacy and PD analyses. Subjects will be analyzed according to the randomized intervention.
Safety	The Safety Population will include all subjects who receive at least one dose of study drug (OCA or placebo). Treatment assignment will be based on the actual treatment that is received.
PK	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.
PD	The PD population for each PD marker will consist of all subjects who receive OCA, have an analyzable sample at baseline and at least one analyzable sample on-treatment, an no major protocol deviations that may potentially affect the PD analysis.

6. GENERAL STATISTICAL METHODS

6.1. Hypothesis and Decision rules

6.1.1. Primary Hypothesis

Null: The time to first occurrence of any of the adjudicated events is equal between placebo and OCA with established local standard of care.

Alternative: The time to first occurrence of any of the adjudicated events are different between placebo and OCA with established local standard of care.

6.1.2. Secondary Hypotheses

Null: the median liver stiffness is equal between placebo and OCA with established local standard of care.

6.2. General Methods

Results will be summarized separately by Cohort, i.e. subjects aged < 2 years (only if supported by results from the first interim analysis) and subjects at least 2 years of age. All endpoints will be analyzed by treatment group.

6.2.1. Analysis for Binary Endpoints

The number and percentage of subjects meeting the binary endpoint will be summarized by timepoint. A CMH test will compare placebo and OCA, adjusting for age group (using median age as cut-off value). A graphical presentation of the proportion of subjects and p-value for difference in proportion over time will be provided as well. Missing values will be imputed as non-responders (NRI).

The above analyses will be performed including only those subjects with available data per timepoint (observed case (OC) approach) as a sensitivity analysis.

6.2.2. Analysis for Continuous Endpoints

Continuous endpoints will be summarized by the following summary statistics (mean (SE), SD, Median, interquartile range (IQR) and minimum and maximum values).

GGT, direct bilirubin, total bilirubin, bile acids (See Section 15) and bile acid totals and plasma levels of fat soluble vitamins D and K.

A mixed model repeated measures (MMRM) will be used to describe the evolution over time in the above continuous endpoints. The model will include terms for randomized treatment group, visit, visit by treatment interaction, age category (using median as the cut-off), and baseline value as covariates. Visits through Week 64 will be included. An unstructured covariance structure will be used. If the computational algorithm fails to converge, the following structures will be tested in the order as specified: Toeplitz, AR(1), and Compound symmetry. If the model still does not converge, age category will be dropped. The corresponding least-square means (LS means), the associated standard errors, the difference in LS Means, the associated 95% confidence interval (CI) and p-value will be presented.

Liver stiffness and markers of FXR activation

The above continuous endpoints will be analyzed using analyzed using Wilcoxon rank sum test. The median differences between treatment groups will be constructed using the Hodges-Lehman estimate. If the median differences between treatment are not approximately symmetrical around the median, the Sign Test may be used.

6.2.3. Analysis for Time-to-Event Endpoints

The time to an event will be calculated in days as the date of first occurrence of the event minus the date of the first intake of study drug plus 1.

If a subject experiences more than one event in a composite endpoint, only the event that occurs first will be included in the analysis.

The log-rank test will be used for comparing treatments adjusted for age (using the median as cut-off). Hazard ratios and the associated 2-sided 5% CIs are estimated by Cox proportional

hazards models adjusted for age group (using median age as cut-off), including treatment as a covariate to estimate the magnitude of the effect. Proportionality of hazards will be assessed using Schoenfeld residuals. The number and proportion of subjects censored and with events will be presented as well.

The median, quartiles, and probabilities of an event at particular points in time will be estimated by the Kaplan-Meier (KM) method. Confidence intervals for medians and quartiles will be based on the Brookmeyer-Crowley method. Confidence intervals for the estimated probability of an event at a particular time point will be generated using the Greenwood formula. KM tabulations will present time to event in days, and KM graphs will present time to event in months along with the number of subjects at risk over time.

6.3. Methods to Manage Missing Data

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

6.3.1. Analysis for Time-to-Event Endpoints

Subjects with no data after randomization will be censored on day 1 (first day of study drug).

When using the Treatment Policy approach, subjects who do not experience an event will be censored at the time of their last contact. Last contact is the date of discontinuation from regularly scheduled study visits.

For the sensitivity analysis excluding data following the intercurrent events (prohibited medication and treatment discontinuation), data will be censored at the time of the intercurrent event or date of last contact, whichever comes first.

If, at a given timepoint, some components of the MELD-NA/PELD scores are available but others are missing, the other components will be carried forward from the most recent measure within 6 months prior to the timepoint (LOCF). If no measure is available within 6 months prior to the timepoint, then the missing component will be imputed with the baseline value.

6.3.2. Analysis for Binary Endpoints

For the primary Estimand of binary endpoint, missing data at a specific timepoint will be considered a non-response for that timepoint.

6.3.3. Analysis for Continuous Endpoints

No imputation for missing data will be performed.

6.3.4. Analysis for Categorical Endpoints

No imputation for missing data will be performed.

7. SUBJECT DISPOSITION

Subject disposition will be tabulated by treatment group and overall and will include the number and proportion of subjects:

- Enrolled
- Randomized (ITT)
- Randomized but not treated
- Safety population
- Evaluable population (mITT)
- PK population
- PD population
- who completed study treatment
- with treatment discontinuation plus reasons for treatment discontinuation
- who completed the study
- who discontinued the study plus reasons for study discontinuation

Proportions will be out of the number of randomized subjects. For the number of subjects enrolled, only the number will be provided, no percentage will be calculated.

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

8. PROTOCOL DEVIATIONS

The investigator is not permitted to deviate from the protocol in any significant way without proper notification to the Sponsor (or designee) unless the Investigator and/or IRB/IEC considers a subject's safety to be compromised if immediate action is not taken. Only the Sponsor may amend a protocol. Any change in study conduct considered necessary by the investigator will be made only after consultation with the Sponsor, who will then issue a formal protocol amendment to implement the change and obtain regulatory approval, as necessary.

Protocol deviations will be listed, and major protocol deviations will be summarized.

No per protocol analyses will be performed.

9. DEMOGRAPHICS AND BASELINE DISEASE CHARACTERISTICS

Demographic variables and baseline disease characteristics will be summarized and presented by treatment group and overall for each the analysis populations.

Demographic variables:

- Age (years) (use months if available for children <2 years of age)
- Age categories 2-18 years cohort
- Age categories <2 years cohort

- Age categories using median age
- Gender
- Race and ethnicity
- Z-score for body height
- Z-score for body weight
- Baseline BMI
- Nutritional score (for subjects 3 months till 5 years of age only)

Baseline disease characteristics:

- Years since diagnosis of biliary atresia
- Age at Kasai (only if this data is available in months)
- Child-Pugh (grade A, B or C)
- MELD-NA score (for subjects >=12 years)
- PELD score (for subjects <12 years)
- Ultrasound elastography of liver stiffness
- Baseline labs (alkaline phosphatase, eGFR, INR, platelets, albumin, GGT, total bilirubin and bilirubin categories (<=ULN; >ULN), direct bilirubin and direct bilirubin categories (<=ULN; >ULN), AST, ALT, total plasma bile acids)
- Aspartate aminotransferase to platelet ratio (APRI)
- UDCA (yes/no)
  - 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 July 1 will be used for the month and day.
  - eGFR is calculated using the following formula:

    eGFR = 0.413 x (height/serum creatinine)

    Where height is in cm and creatinine is in mg/dL.

For categorical variables, the number and percentage of subjects within each category, including a category for missing data, will be presented. For continuous variables, summary statistics will include the number of subjects and the mean, standard deviation (SD), standard error of the mean (SEM), median, 25^{th} and 75^{th} quartiles, minimum, and maximum values.

The above analyses will be performed on both the ITT and safety populations, unless the safety population is the same as the ITT population.

10. PRIOR THERAPY AND MEDICAL HISTORY

Prior therapy and medical history are any therapy or diseases that occurred, or any medication started prior to the first day of study drug dosing or prior to randomization date if subjects were never dosed.

Prior therapy will be mapped to Anatomic Therapeutic Chemical (ATC) class and preferred term (PT) using WHO Drug Dictionary, September 2024 and summarized by ATC class, preferred term, and treatment in the Safety population. Summaries will be ordered by descending order of incidence of ATC class and preferred term within each ATC class.

Medical history will be mapped to preferred terms and system organ classes using MedDRA dictionary and summarized by system organ class, preferred term, and treatment group in the Safety population. Summaries will be ordered by descending order of incidence of system organ class (SOC) and PT within each system organ class.

11. EFFICACY ANALYSES

The efficacy endpoints of the study are defined below. Additionally, the planned analyses for these endpoints are described. Statistically significant treatment differences will be declared if the resulting p-value is 0.05 or less. All tests will be two-sided.

11.1. Primary [Efficacy] Endpoint

The primary efficacy variable is time to the first occurrence of any component of the composite clinical outcome.

- Death (all-cause)
- Liver transplant
- Pediatric end-stage liver disease (PELD) ≥17/MELD-NA ≥15
- Hospitalization (as defined by a stay of 24 hours or greater) for new onset or recurrence of:
  - Variceal bleed
  - Hepatic encephalopathy (as defined by a West Haven score of ≥2)
  - Spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis)
- Clinically evident ascites related to portal hypertension (diuretic-resistant ascites requiring therapeutic paracentesis at a frequency of at least twice in a month)

The hypothesis for testing the primary endpoint is described in Section 6.1.1.

11.2. Analyses for Time-to-Event Endpoints

11.2.1. Primary Analysis

The primary efficacy endpoint will be conducted using the ITT population (see Section 5). Only adjudicated events will be included in the analysis.

If a subject experiences more than one event in the composite primary endpoint, only the event that occurs first will be included in the analysis.

The primary estimand and method of analysis are described in Section 3.2.1 and handling of missing data is described in Section 6.3.1.

A listing for time to adjudicated event will be provided.

11.2.2. Sensitivity/Supplementary Analyses

The following sensitivity/supplementary analyses will be carried out for the primary endpoint.

Table 5: Supplemental Estimands to the Primary Estimand

Estimand	Population	Endpoint	Intercurrent Events	Strategy for Intercurrent Events	Population-Level Summary
Sensitivity/Supplementary Analysis to Primary Estimand					
1	mITT	See primary estimand	Same as primary estimand		See primary estimand
2	See primary estimand	See primary estimand	Same IC as primary estimand Data will be censored at the time of the first IC.		See primary estimand

In addition, a tipping point analysis using multiple imputation (MI) will be performed to investigate departures from ignorable censoring as described in Appendix C.

11.3. Secondary [Efficacy] Endpoints

All secondary efficacy endpoints will be analyzed based on the ITT Population including all data, regardless of ICs (i.e., the use of prohibited medication and treatment discontinuation).

Additionally, liver stiffness by transient elastography will be analyzed based on the mITT population.

11.3.1. Composite Endpoint of Biomarkers

The composite endpoint of biomarkers – proportion of subjects with improvement in GGT and direct (conjugated) bilirubin. A subject is a responder if:

1. Percent change from baseline in GGT is ≤ -40%
2. Percent change from baseline in bilirubin is ≤ -25%
3. The month-4 assessments will be used for the first interim analysis, and final assessments used for the EOS analysis

The proportion of subjects who are biochemical responders will be summarized and analyzed as outlined in Section 6.2.1. Handling of missing data is detailed in Section 6.3.2.

11.3.2. Biomarkers of Hepatobiliary Function

Summary statistics (Section 6.2.2) will be provided for GGT, direct and total bilirubin value, change from baseline and percent change from baseline. Change from baseline and percent change from baseline will be analyzed as described in Section 6.2.2. Mean (+/- SE) and LS Mean (+/- SE) change from baseline will be plotted over time by treatment arm.

11.3.3. Plasma Fat Soluble Vitamins D and K

Plasma fat soluble vitamins will be analyzed similarly as biomarkers for hepatobiliary function.

Summary statistics (Section 6.2.2) will be provided for Plasma fat soluble vitamins D and K value, change from baseline and percent change from baseline. Change from baseline will be analyzed as described in Section 6.2.2. Mean (+/- SE) and LS Mean (+/- SE) change from baseline will be plotted over time by treatment arm.

11.3.4. Transient Elastography (TE)

Transient elastography will be measured using the Fibroscan® TE device at selected sites where the device is available. Liver stiffness by TE is measured in kiloPascale (kPa).

Summary statistics (Section 6.2.2) will be provided for liver stiffness and change from baseline. Change from baseline in liver stiffness will be analyzed as described in Section 6.2.2. Means (+/- SE) liver stiffness and median (IQR) liver stiffness and change from baseline in liver stiffness will be plotted over time by treatment arm. In addition, a listing of TE measurements of liver stiffness will be provided.

11.4. Exploratory [Efficacy] Endpoints

Acceptability and swallowability will be evaluated using 5 categories (Swallowed and/or consumed; partially swallowed and/or consumed; spat out the majority of the dose regimen; inhaled or coughed during deglutition; infused to take) during screening.

Palatability will be assessed using a four-point criterium for children < 4 years of age, and a visual 5-point hedonic scale with a correlated 100-point visual analogue scale (VAS) at Week 6.

Acceptability and swallowability, and Palatability will be summarized by tablet formulation and treatment group.

For categorical questions (swallowability, palatability using a 4-point scale, and palatability using the 5 hedonic subscales, the number and percentage of subjects per category will be presented. Only non-missing data (OC approach) will be summarized.

For the VAS part, summary statistics will be provided as described in Section 6.2.2.

A by age-group subgroup (using median age) summary will also be presented.

11.5. Baseline Variables

The baseline value for statistical analyses of quantitative parameters is defined as the mean of all available study evaluations after the subject signs informed consent and prior to or on the day of first administration of study drug, unless otherwise specified. If there is only one evaluation prior

to the first administration of study drug then the available data from this evaluation will be used as the baseline value.

The baseline value for analyses of lipid parameters is defined as the last fasted evaluation prior to the first administration of study drug.

The baseline value for analyses of quantitative parameters (e.g. normal/abnormal) is defined as the last evaluation prior to the first administration of study drug.

11.6. Multiple Comparisons/Multiplicity

In addition to the final data analyses, the IA based on the biochemical endpoints will be tested at the month 4 data cut to detect early signs of efficacy. The second IA will be performed based on the primary endpoint, the composite score of the clinical outcomes. The trial can be terminated for futility at either IA. No adjustment for alpha will be performed. The 2-sided alpha allocated to all testing for the final data analyses will be 0.05.

For the final analysis, the hypothesis testing of the key secondary endpoint will compare placebo and OCA provided that the primary efficacy endpoint comparison is statistically significant. If the primary comparison is not statistically significant then comparison of the key secondary endpoint will be considered as supportive or supplemental. All hypotheses will be tested at the nominal 0.05 alpha level.

11.7. Subgroup Analyses of Efficacy

The primary and selected secondary efficacy endpoints (liver stiffness, GGT and C4) will be analyzed by subject subgroups based on the ITT population. Subgroups will be assessed at baseline. Sensitivity analyses will not be run on subgroup analyses. Subgroup analyses will only be run if there is a sufficient number of subjects in each group (>10 subjects per group).

Baseline subgroups of interest are as follows:

- Years since Kasia procedure
- GGT: >ULN, <= ULN
- Total bilirubin: >ULN, <= ULN

12. SAFETY ANALYSES

Safety analyses include exposure to study treatment, treatment-emergent adverse events (TEAEs), clinical labs, electrocardiograms (ECGs), vital signs, and any abnormal findings observed during physical examinations after first dose and through discontinuation of study treatment + 30 days. Safety analyses will be performed using the Safety population including by treatment groups. No inferential comparison of safety endpoints will be performed, unless otherwise specified.

12.1. Exposure to Study Treatment

The extent of exposure will be summarized using descriptive statistics and by duration categories.

Duration of exposure (days) will be calculated as follows:

- Exposure to study drug = Date of last dose – Date of first dose +1
- Duration of each incidence of temporary study drug interruption will be calculated as date of restart of study drug – Date of temporary interruption of study drug + 1. The total duration of temporary study drug interruption is the sum of the durations of temporary interruptions of study drug per subject.
- Duration of exposure = exposure to study drug – duration of temporary interruption

The cumulative dose will be calculated by adding the doses taken by a single subject during the study and will be calculated using descriptive statistics.

A summary of the number and percentage of subjects who had missed doses, interruptions, and premature discontinuations will be provided.

Subject's overall compliance will be summarized using descriptive statistics and by compliance categories. Subject's overall compliance (%) with study drug will be calculated as follows:

100 x (number of tablets consumed during study / number of tablets expected to be consumed during study)

where

Number of tablets expected to be consumed during study = number of dispensed tablets during the study

and

Number of tablets consumed during the study = number of tables dispensed – number of tablets returned

12.2. Adverse Events

Per ICH E6 (R2), an adverse event (AE) is any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to study drug.

AEs include but are not limited to: (1) a worsening or change in nature, severity or frequency of condition(s) present at the start of the study; (2) subject deterioration due to primary illness; (3) intercurrent illness; and (4) drug interaction. For reporting purposes, pregnancy is not considered an AE.

AEs will be mapped to preferred terms (PTs) and system organ classes (SOCs) using MedDRA dictionary and graded for severity according to CTCAE version 5.0. Subjects experiencing the same event more than once will be counted only once at the most severe grade and the closest relationship to study treatment.

12.2.1. Treatment-emergent Adverse Events

A TEAE is any event not present before the initiation of study drug or any event already present that worsens in either intensity or frequency following exposure to study drug.

Following AEs are considered treatment-related:

- An AE starting on or after the first study drug dose and within 30 days after the last dose of study drug.
- An AE occurring prior to the first study drug dose that worsens (increase in grade) after the first study drug dose.
- An AE starting more than 30 days after the last dose of study drug with a relationship to study drug of "Possible", "Probable", or "Definite" as assessed by the investigator.
- An AE with a missing start date but with an end date after the first study drug dose date.
- If it cannot be determined whether the AE is treatment-emergent due to a partial start date, then it will be considered as such.

12.2.2. Serious Adverse Events

An AE is considered "serious" if, in the view of either the investigator or Sponsor, it results in any of the following outcomes:

- Death,
- Is immediately life-threatening,
- Requires in-patient hospitalization or prolongation of existing hospitalization,
- Results in persistent or significant disability/incapacity,
- Results in a congenital anomaly/birth defect, or
- Is an important medical event that may jeopardize the subject or may require medical intervention to prevent one of the outcomes listed above. Examples of such events are intensive treatment in an intensity room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization; or development of drug dependency or drug abuse. Any malignancy is to be reported as a serious AE.

AEs not considered to be a serious AE (SAE) are hospitalizations for:

- Routine monitoring of the studied indication and not associated with any deterioration in condition or AE.
- Elective treatment for a pre-existing condition that did not worsen such as elective cosmetic surgery or medical procedure
- Respite care or observation when there is no AE associated with the hospitalization

12.2.3. Adverse Events of Interest

The following AEs of Interest (AEI) will be summarized:

- Pruritus
    - Defined as any preferred term withing the Pruritis Not Elsewhere Classified (NEC) high level term or any preferred term including pruritus
- Hepatic disorders
- 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
- Cholecystitis/cholelithiasis
    - Defined by Gallbladder related disorders narrow SMQ and Gallstone related disorders narrow SMQ
- Renal
    - Defined as events included in the following broad SMQs: acute renal failure, chronic kidney disease, proteinuria, renovascular disorders, or tubulointerstitial diseases
- Cardiovascular
    - Defined by the Embolic and thrombotic events broad SMQ, Ischemic heart disease broad SMQ, or Central nervous system vascular disorders narrow SMQ
    - These events may be reassessed stratified by statin (or other lipid lowering medication) use
- Dyslipidemia
    - Defined by Dyslipidemia SMQ
    - These events may be reassessed stratified by stratified by statin (or other lipid lowering medication) use

12.2.4. Treatment-emergent Adverse Events Analyses

The counts and percentages of subjects experiencing any TEAE and TEAEs by PTs and SOCs in the following categories will be presented by study treatment arm:

- TEAEs
- TEAEs by highest severity (mild, moderate, severe): 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 will be considered severe for this summary.

- TEAEs by relationship to treatment (related, not-related): Related AEs are those with relationships reported by the investigator as “Definite”, “Probably” or “Possibly”, and unrelated AEs are those with relationship reported as “Unlikely” or “Not related”. .At each level of subject summarization, a subject is classified according to the closest relationship if the subject reported 1 or more events. AEs with missing relationship will be considered related for this summary.
- SAEs
- Related TEAEs: this is the subset of AEs reported by the investigator as “possibly”, “probably” or “definitely” related to study therapy.
- Severe TEAEs: this is the subset of AEs which are at least grade 3 in severity
- AEIs
- AEIs by highest severity (mild, moderate, severe): 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 will be considered severe for this summary.
- TEAEs Leading to Study Treatment Discontinuation: this is a subset of the AEs where action taken with study treatment is checked as “Drug withdrawn” or where “Subject discontinued from study” is checked.
- TEAEs Leading to Study Treatment Interruption: this is a subset of the AEs where action taken with study treatment is checked as “Drug interrupted”.
- Deaths: this is the subset of AEs where “fatal” is checked
- Subject incidence of TEAEs by MedDRA SOC, PT, and closest relationship to study drug (Related/Not Related)

Summaries will be presented by PT, sorted in order of decreasing frequency in the treatment arm.

12.2.5. Missing or Partial Date Handling

If only a partial date is available and is required for calculation, the following standards will be applied:

- Diagnostic Date (eg, BA diagnostic 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).
- Start Dates (eg, event date, AE onset date, or start date of medication)
  - For missing start day only – Day will be imputed as the first day of the month (ie, 1) with the following exception: if the partial date falls in the same month and year as the date being used in the calculation (eg, first dose date, informed consent date), then the partial date will be imputed to equal the date being used for the calculation.

  - For missing start day and month – Day and month will be imputed as the first day of the year (ie, 1 January) with the following exception: if the partial date falls in the same year as the date being used in the calculation (eg, first dose date, informed consent date), then the partial date will be imputed to equal the date being used for the calculation.
  - Imputed start dates must be prior to the stop date. Otherwise, the start date will be further imputed to be the same as the stop date.
  - *Additional rules can be added.*
- Stop Dates (eg, AE resolution date or stop date of medication)
  - 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, 31 December).
  - Imputed stop dates must be on or after the start date. Otherwise, the stop date will be further imputed to be the same as the start date.

12.3. Clinical Laboratory Evaluations

Laboratory parameters will be summarized in both conventional units and the standard international (SI) system of units. Only data from central lab will be summarized.

Laboratory parameters will be summarized using descriptive statistics (see Section 6.2.2) at baseline and at each scheduled timepoint (see Appendix B). The change from baseline will also be summarized. Rules for windowing are described in Appendix B. In addition, graphs will be created for GGT, direct bilirubin, AST, ALT and alkaline phosphatase.

Change from baseline at a particular time point will be computed by subtracting the baseline value from the post-baseline value.

Derivation of the baseline value is described in Section 11.5. The baseline value for analyses of lipid parameters is defined as the last fasted evaluation prior to the first administration of study drug.

In addition, shift tables from baseline based on normal ranges will be provided for hematology, coagulation and chemistry to assess changes from baseline to each pre-planned post-baseline time point. For each timepoint, the worse value will be selected. For all summaries of laboratory abnormalities, the denominator is the number of subjects with non-missing postbaseline values per analysis timepoint.

In addition, the number and percentage of subjects with liver toxicity will be summarized for reach of the below categories:

- AST > 3 times, 5 times, 10 times and 20 times the upper limit of normal (ULN)
- ALT > 3 times, 5 times, 10 times and 20 times the ULN
- AST or ALT > 3 times the ULN and total bilirubin > 2 times the ULN

The baseline value for analyses of quantitative parameters (e.g. normal/abnormal) is defined as the last evaluation prior to the first administration of study drug.

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.

For tabulations, the unscheduled post-baseline values will generally be excluded from summary tables but will be included in data listings. Unscheduled visits will be considered for analyses of shift from baseline to worst value (low-normal-high).

A selection of clinical laboratory data will be presented in by-subject data listings.

12.3.1. Hematology

Hematology includes hemoglobin, hematocrit, white blood count with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils), platelets, red blood cell count (including MCV, MCH, MCHC).

12.3.2. Serum Chemistry

Serum chemistry includes albumin, BUN, creatinine, conjugated (direct) bilirubin, total bilirubin, ALT, AST, ALP, GGT, creatinine phosphokinase, electrolytes (calcium, chloride, magnesium, phosphorus, potassium, sodium), glucose, total protein, bicarbonate, free fatty acids and blood lipids (total cholesterol, LDL, HDL, VLDL fractions and triglycerids).

Adjudicated DILI results will be summarized.

12.3.3. Other Laboratory Parameters

Other laboratory lab parameters include coagulation assessments (PT, PTT and INR), glycemic control measures (fasting plasma glucose measures), thyroid function (T3, T4 and TSH), vitamin B and K levels, PD assessments (C4, FGF-19 and analysis of conjugated and unconjugated bile acids) and aspartate aminotransferase to platelet ratio (APRI).

12.4. Anthropomorphic Measurements and Vital Signs

The results and change from baseline to each on-study evaluation visit will be summarized for height, weight, z-score, triceps skinfold thickness, mid-arm circumference measurement, systolic and diastolic blood pressure, sitting heart rate, respiratory rate and body temperature.

The abnormal findings observed during physical examinations will be summarized for each study visit.

12.5. Electrocardiograms (ECGs)

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 >440 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.

12.6. Prior and Concomitant Medications

Prior medications are medications that were started prior to the first dose of study drug, and concomitant medications are medications that were started after the initial dose of study drug and. Prior and concomitant medications will be mapped to Anatomic Therapeutic Chemical (ATC) class and preferred term using WHO Drug Dictionary version B3, September 2024 and summarized by ATC class, preferred term, and treatment in the Safety population. Summaries will be ordered by descending order of incidence of ATC class and preferred term within each ATC class.

12.7. Subgroup Analyses of Safety

No subgroup analyses will be performed for safety endpoints.

13. INTERIM ANALYSES

13.1. Introduction

Two interim analyses will be performed.

A first interim analysis (IA) will be performed after approximately 25 to 30 subjects reach Month 4. Assessments of GGT, total and direct (conjugated) bilirubin, total plasma bile acids and bile acid totals and PK will be evaluated by a DSMB. The DSMB is an independent committee that includes individuals who will not be involved in the study and is described in the DMC Charter. The DSMB may terminate the study for futility and will advise on the potential inclusion of subjects <2 years of age into the study as a separate cohort.

Responders are defined as having a ≥40 reduction from baseline in GGT and a ≥25% reduction from baseline in direct bilirubin at Month 4.

Assuming a 10% and 55% response rate for the placebo and OCA groups, respectively, and a total sample size of 30, the mean (95% CI) for the treatment difference is 45% (16% to 74%). The treatment difference of at least 15% is considered a clinically important difference. The study will be terminated for futility if the futility boundary of 15% is crossed (i.e., if the treatment difference falls outside of the lower bound of the 95% CI).

A second IA will be performed based on the primary endpoint (time to first occurrence of clinical outcomes). The timing of the second IA is approximately 12 months after the first interim analysis, but the exact timing is to be confirmed by the DSMB. A group sequential design and conditional power approach will be utilized (Jennison and Turnbull, 2000). At the time of the second IA, the interim data will be analyzed to determine the conditional power. If the power is lower than 80%, then the trial will be terminated for futility.

13.2. Interim Analyses and Summaries

For the first interim analysis the following will be presented as a minimum:

- Demographics and baseline disease characteristics
- Number and percentages of subjects with AEs as listed under Section 12.2.4.
- Number and percentage of subjects with TEAEs by PT.
- Selected laboratory parameters over time including total bilirubin and GGT. These analyses may be stratified by age quartiles
- PELD/MELD-NA over time

For the second interim analysis, at a minimum, the conditional power for the primary endpoint will be calculated according to Jennison and Turnbull 2000, p 205-220. If the conditional power is lower than 80% the study will be stopped for futility. Details of the calculations can be found in Appendix C.

14. PHARMACOKINETIC ANALYSES

Concentrations of total OCA will be expressed in ng-OCA equivalents/mL but referred to as ng/mL total OCA to simplify presentation. Data for unconjugated OCA, glyco-OCA and tauro-OCA will be received from the bioanalytical laboratory in units of ng/mL and will be converted to units of µM. The conversion formula is: analyte concentration in µM = analyte concentration in ng/mL / analyte molecular weight. Molecular weights are produced in Table 6. Total OCA in ng/mL is calculated as (unconjugated OCA [µM] + glyco-OCA [µM] + tauro-OCA [µM]) / molecular weight for unconjugated OCA.

Table 6: Bile Acid Totals and Individual Bile Acids with Molecular Weights

Analyte	Individual Bile Acids (Molecular Weight)
Unconjugated OCA	420.6
Glyco-OCA	477.7
Tauro-OCA	527.8

If any component of total OCA is missing, total OCA will not be calculated. If a sample for any component of total OCA is successfully analyzed and the concentration reported as below the lower limit of quantitation (BLQ), the value will be imputed to 0.

Concentrations of unconjugated OCA, glycol-OCA, tauro-OCA and total OCA will be listed by subject and visit and summarized by adult equivalent dose and visit. Summary statistics will include n, arithmetic mean, SD, coefficient of variation (CV%), geometric mean, geometric coefficient of variation (GCV%), median, minimum and maximum. Concentration data from this study may also be included in cross-study analyses, which will be reported separately.

15. PHARMACODYNAMIC ANALYSES

Concentrations of markers for FXR activation (C4 and FGF-19, total plasma bile acids and plasma bile acid totals) will be summarized as described in Section 6.2.2. Mean (+/- SE) and median (IQR) will be plotted over time by treatment arm. Individual bile acids will be listed.

Bile acid data will be received from the laboratory in units of ng/mL and will be converted to units of µM for analysis and presentation. Unit conversion is bile acid concentration (µM) = bile acid concentration (ng/mL) / bile acid molecular weight. Table 7 provides molecular weights for the 15 individual bile acids. Bile acid total is the sum of the individual unconjugated, glycine-conjugated and taurine-conjugated forms of a bile acid. Total bile acids are the sum of all 15 individual bile acids. If any component of a calculated total is missing, the calculated total will also be reported as missing. If any component of a calculated total is BLQ, the calculated value will be reported as BLQ in individual listings and imputed as 0 for statistical summaries and analyses.

Table 7: Bile Acid Totals and Individual Bile Acids with Molecular Weights

Bile Acid Totals	Individual Bile Acids (Molecular Weight)		
Total UDCA	Glyco-UDCA (449.6 g/mol)	Tauro-UDCA (499.7 g/mol)	UDCA (392.6 g/mol)
Total CDCA	Glyco-CDCA (449.6 g/mol)	Tauro-CDCA (499.7 g/mol)	CDCA (392.6 g/mol)
Total DCA	Glyco-DCA (449.6 g/mol)	Tauro-DCA (499.7 g/mol)	DCA (392.6 g/mol)
Total CA	Glyco-CA (465.6 g/mol)	Tauro-CA (515.7 g/mol)	CA (408.6 g/mol)
Total LCA	Glyco-LCA (433.6 g/mol)	Tauro-LCA (483.7 g/mol)	LCA (376.6 g/mol)

Additional pharmacodynamic analyses, if any, will be specified in a separate report.

16. REFERENCES

ICH E9: Statistical Principles for Clinical Trials, EMA/CHMP/ICH Committee for Medicinal Products for Human Use, 24 August 2018.

ICH E9 (R1): Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials, EMA/CHMP/ICH Committee for Medicinal Products for Human Use, 17 February 2020.

ICH E6 (R2): Good Clinical Practice: Integrated Addendum to ICH E6(R1). 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), March 2018.

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.

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

Jennison and Turnbull (2000), p205-220. Groups Sequential Methods Applications to Clinical Trials. Chapman & Hall/CRC.

Lipkovic I, Ratitch B and O'Kelly Michael. Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints. Pharmaceutical Statistics 2016; 15: 216-229.

Thompson C, Lombardi D, Sjostedt P et Al. Best practice recommendations regarding the assessment of palatability and swallowability in the development of oral doses for pediatric patients. Therapeutic Innovation & Regulatory Science. 2015; 49(5): 647-58.

APPENDIX A. DATA PRESENTATION CONVENTIONS:

For categorical variables, the number and percentage of subjects within each category of the parameter, including a category for missing data, will be presented. Percentage calculations will be based on the analysis population, unless otherwise specified. Percentages are rounded to 1 decimal place, unless otherwise specified. For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness. For example, if none of the subjects discontinue due to "lost to follow-up," this reason will be included in the table with a count of 0. Percentages based on frequency counts will be presented as a whole number (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%".

For continuous variables, summary statistics will include the number of subjects and the mean, standard deviation (SD), standard error of the mean (SEM), median, 25th and 75th quartiles, minimum, and maximum values. The precision of summary statistics, unless otherwise specified, will be as follows: mean and median to 1 more decimal place than the raw data, and SD and SEM to 2 decimal places more than the raw data. In general, the decimal places should not exceed 3 decimal places unless appropriate. Confidence intervals (CIs) will be provided and will be rounded to 1 decimal place, unless otherwise specified, in the table and listing shell.

All statistical analyses will be performed using SAS statistical software Version [4.1] or higher unless otherwise noted.

APPENDIX B. DEFINITION AND USE OF VISIT WINDOW IN REPORTING

Duration variables 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 365.25/12 (approximately 30.4).
- 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)/365.25, where date1 ≥first dose date duration in years = (date2 – date1)/365.25, where date1<first dose date.

Analysis windows will be defined around the target date of the visit planned in the protocol. Day 1 is the first day of receiving study drug (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 one day less than the lower bound of the next interval.

If more than 1 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. The below tables show the analyses windows at each timepoint for the different types of endpoints.

Table 8: Analysis Visits

	Serum chemistry, hematology, coagulation, lipoprotein, cholesterol, vital signs and body weight	Plasma FGF-19, C4 and endogenous bile acids	Fat soluble vitamins D and K	Thyroid function (T3, T4, TSH) and Height	Z-score	Triceps Skinfold	MELD-NA /PELD and Child Pugh Score
Screening	< Day 1	< Day 1	< Day 1	< Day 1	< Day 1	< Day 1	< Day 1
Day 1	Day 1	Day 1	Day 1	Day 1	Day 1	Day 1	Day 1
Week 2	14 (2, 21)	14 (2, 21)	14 (2, 21)				14 (2, 21)
Week 4	28 (22, 35)	28 (22, 35)	28 (22, 63)			28 (2, 49)	28 (22, 35)
Week 6	42 (36, 56)	42 (36, 56)			42 (2, 84)		42 (36, 56)
Week 10	70 (57, 84)	70 (57, 98)				70 (50, 98)	70 (57, 84)
Week 14	98 (85, 112)		98 (64, 175)				98 (85, 154)
Week 18	126 (113, 147)	126 (99, 168)			126 (85, 189)	126 (99, 168)	
Week 24	168 (148, 189)						
Week 30	210 (190, 231)	210 (169, 329)		210 (2, 329)		210 (169, 252)	210 (155, 231)
Week 36	252 (232, 273)		252 (176, 294)		252 (190, 294)		252 (232, 273)
Week 42	294 (274, 315)					294 (253, 336)	294 (274, 371)
Week 48	336 (316, 357)		336 (295, 392)		336 (295, 378)		
Week 54	378 (358, 399)					378 (337, 413)	
Week 60	420 (400, 434)				420 (379, 434)		
Week 64	448 (435, +∞)	448 (330, +∞)	448 (393, +∞)	448 (330, +∞)	448 (435, +∞)	448 (414, +∞)	448 (372, +∞)

Table 8: Analysis Visits (Continued)

	Palatability	**Ultrasound Elastography of Liver Stiffness**	**FIB4 Score**	**APRI and ECG**			
Screening		< Day 1	< Day 1	< Day 1			
Day 1		Day 1	Day 1	Day 1			
Week 2			14 (2, 21)				
Week 4			28 (22, 35)				
Week 6	42 (2, +∞)		42 (36, 56)				
Week 10			70 (57, 84)				
Week 14			98 (85, 112)				
Week 18		126 (2, 189)	126 (113, 147)				
Week 24			168 (148, 189)				
Week 30			210 (190, 231)	210 (2, +∞)			
Week 36		252 (190, 315)	252 (232, 273)				
Week 42			294 (274, 371)				
Week 48							
Week 54		378 (316, 413)					
Week 60							
Week 64		448 (414, +∞)	448 (372, +∞)				

APPENDIX C. STATISTICAL METHODOLOGY DETAILS