OBETICHOLIC ACID (OCA)

Protocol 747-213

A Phase 2, Double-Blind, Randomized, Parallel Group Study Evaluating the Efficacy, Safety, and Tolerability of Obeticholic Acid Administered in Combination with Bezafibrate in Subjects with Primary Biliary Cholangitis who had an Inadequate Response or who were Unable to Tolerate Ursodeoxycholic Acid

Statistical Analysis Plan

Author: [redacted]

[redacted]Biostatistics and Data Management

Sponsor: Intercept Pharmaceuticals, Inc

305 Madison Avenue

Morristown, NJ 07960

Protocol Version and Date: Version 7.0, 08 Oct 2024

Analysis Plan Version: 4.0

Analysis Plan Date 15 Oct 2025

CONFIDENTIAL

The information contained herein is the property of Intercept Pharmaceuticals, Inc. and may not be reproduced, published, or disclosed to others without written authorization of Intercept Pharmaceuticals, Inc.

APPROVAL

Upon review of this document, including table and listing shells, the undersigned approves the Statistical Analysis Plan. The analysis methods and data presentation are acceptable.

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 13, 2025 09:36:09 EDT

13-Oct-2025
Date

[redacted], Biostatistics
Intercept Pharmaceuticals, Inc.

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 13, 2025 13:46:09 EDT

13-Oct-2025
Date

[redacted] Biostatistics
Intercept Pharmaceuticals, Inc.

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 15, 2025 08:41:46 CDT

15-Oct-2025
Date

[redacted] Clinical Research
Intercept Pharmaceuticals, Inc.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS 6
1. SCOPE 9
2. STUDY OBJECTIVES AND HYPOTHESES 9
2.1. Study Objectives 9
2.1.1. Primary Objective 9
2.1.2. Secondary Objectives 9
2.1.3. Additional Objectives 9
2.2. Hypotheses 10
2.2.1. Primary Hypothesis 10
3. SUMMARY OF STUDY DESIGN 10
3.1. Overall Study Design 10
3.2. Study Design Diagram 11
3.3. Randomization and Blinding 12
3.4. Sample Size Determination 13
4. ANALYSIS POPULATIONS AND APPROACHES TO ANALYSIS 13
4.1. Intent-to-Treat Population 13
4.2. Modified Intent-to-Treat Population 13
4.3. The Safety Population 13
4.4. PK Population 13
4.5. The LTSE Population 13
5. GENERAL CONSIDERATION 14
5.1. Baseline Definitions 14
5.2. Partial Dates 14
5.3. Data Conventions 15
5.4. Calculations 16
5.5. Analysis Windows 18
6. SUBJECT DISPOSITION 18
6.1. Protocol Deviations 19
7. DEMOGRAPHICS AND BASELINE DISEASE CHARACTERISTICS 19
7.1. Demographics and Baseline Characteristics 19
7.2. Baseline PBC Disease Characteristics 20

8. MEDICAL HISTORY 20
9. EFFICACY ANALYSES 20
9.1. Primary Efficacy Endpoint 21
9.2. Secondary Efficacy Endpoints 22
9.3. Additional Efficacy Endpoints 23
9.3.1. Quantitative Endpoints 24
9.3.2. Prognostic Endpoints 24
9.3.2.1. Change from Baseline in Mayo Risk Score 24
9.3.2.2. Child-Pugh 24
9.4. Analyses of LTSE Efficacy Endpoints 25
9.5. Interim Analyses and Data Monitoring Committee 25
9.5.1. Interim Analyses 25
9.5.2. Data Monitoring Committee 26
9.6. Multiple Comparisons/Multiplicity 26
9.7. Subgroup Analyses of Efficacy 27
10. SAFETY ANALYSES 27
10.1. Exposure to Study Treatment 27
10.2. Adverse Events 27
10.2.1. Treatment-emergent Adverse Events 28
10.2.2. Serious Adverse Events 28
10.2.3. Adverse Events of Special Interest 29
10.2.4. TEAE Analyses 29
10.2.4.1. Pruritus 30
10.2.4.2. Hepatic Injury 31
10.2.4.3. Myalgia 31
10.2.4.4. Renal Impairment 32
10.3. Clinical Laboratory Evaluations 32
10.4. Vital Signs 33
10.5. Electrocardiograms 34
10.6. Physical Exam 34
10.7. Prior and Concomitant Medications 34

[redacted]

10.9. Subgroup Analyses of Safety 35

[redacted]

13. DEVIATIONS FROM PROTOCOL 35
14. REFERENCES 35
APPENDIX A. DATA PRESENTATION CONVENTIONS: 36
APPENDIX B. SUMMARY OF CHANGES: 37

LIST OF TABLES

Table 1: Child-Pugh Score 25

LIST OF FIGURES

Figure 1: Study Design for the Double-Blind/Follow-Up and LTSE Treatment Periods (Study 747-213) 11

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or Specialist Term	Explanation
AE	adverse event
AESIs	Adverse Events of Special Interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
[redacted]	
AR (1)	first-order autoregressive
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
BZF	bezafibrate
[redacted]	
C4	7-alpha-hydroxy-4-cholesten-3-one
[redacted]	
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CP	Child-Pugh
CS	clinically significant
CSR	clinical study report
DAP	data access plan
DB	double-blind
DMC	Data Monitoring Committee
DOB	date of birth
eCRF	electronic case report form
ECG	electrocardiogram
[redacted]	
[redacted]	
GGT	gamma-glutamyl transferase
ICH	International Conference on Harmonisation

Abbreviation or Specialist Term	Explanation
INR	international normalized ratio
ITT	Intent-to-Treat
IWRS	interactive web response system
KM	Kaplan-Meier
LTSE	long term safety extension
LS	least-square
[redacted]	
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified Intent-to-Treat
MRS	Mayo Risk Score
[redacted]	
NCS	not clinically significant
[redacted]	
OCA	obeticholic acid
OPTN	Organ Procurement and Transplantation Network
PBC	primary biliary cholangitis
PD	pharmacodynamic
PK	pharmacokinetic
[redacted]	
PT	preferred term
QD	once daily
QTc	corrected QT interval
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SEM	standard error of the mean
SI	standard international unit
SMQ	Standardized MedDRA Query
SOC	system organ classes

Abbreviation or Specialist Term	Explanation
SE	standard error
[redacted]	
TEAE	treatment-emergent adverse event
UDCA	ursodeoxycholic acid
ULN	upper limit of normal
[redacted]	
WHO	World Health Organization

1. SCOPE

This statistical analysis plan (SAP) describes the statistical analyses and data presentations planned for protocol 747-213 which includes the Double-Blind (DB) period and Long-Term Safety Extension (LTSE) period. It 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 off before the study unblinding of the interim analysis. It could be further amended in the later database locks. 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.

2. STUDY OBJECTIVES AND HYPOTHESES

2.1. Study Objectives

2.1.1. Primary Objective

The primary objective is to assess the effects of the combination of obeticholic acid (OCA) and Bezafibrate (BZF) on ALP in comparison to BZF alone in subjects with Primary Biliary Cholangitis (PBC).

2.1.2. Secondary Objectives

The secondary objectives are to assess the effects of the combination of OCA and BZF in comparison to BZF alone in subjects with PBC on the following:

- Biochemical disease markers, including ALP, GGT, ALT, AST, total and conjugated bilirubin and a lipid panel
- Biomarkers of bile acid synthesis and homeostasis, including 7α-hydroxy-4-cholesten-3-one (C4) and bile acids
- Safety and tolerability

2.1.3. Additional Objectives





2.2. Hypotheses

2.2.1. Primary Hypothesis

The primary efficacy analysis will test the following hypothesis for each comparison between OCA + BZF and BZF (i.e. OCA + BZF 200 mg vs BZF 200 mg and OCA + BZF 400 mg vs BZF 400 mg) with two-sided α of 0.05:

$$H_0: \mu_{ob}=\mu_b$$

$$H_a: \mu_{ob}\neq\mu_b$$

Where

- μ_{ob} represents the mean change of ALP from baseline to Week 12 in the combination of OCA and BZF group
- μ_b represents the mean change of ALP from baseline to Week 12 in the BZF group

3. SUMMARY OF STUDY DESIGN

3.1. Overall Study Design

This is a Phase 2, proof-of-concept, randomized, double-blind (subject and site staff), parallel-group study in subjects with PBC as defined in Section 7 of the protocol.

This study will evaluate the efficacy, safety, and tolerability of OCA (5 mg and 10 mg) administered in combination with 2 different BZF doses (200 mg and 400 mg) or BZF alone (at two doses, 200 mg and 400 mg) in up to 72 subjects with PBC over at least 12 weeks. An interim analysis will be performed when at least 40 subjects have completed the DB period to guide decision-making for the Phase 3 study. No futility or superiority stopping rules will apply for the interim analysis. During the course of the study, both safety and efficacy data will be reviewed by unblinded Sponsor team members who are not involved in the conduct of the study.

Subjects who are randomized to combination groups will receive OCA 5 mg once daily (QD) from Day 1 to the day before the Week 4 Visit, followed by OCA 10 mg QD from the Week 4 Visit through the end of the study. To preserve the study blind, appearance-matched placebo tablets for OCA and/or BZF will be administered to subjects in each treatment group from Day 1 to Week 12 as shown in the study design diagrams for the DB and LTSE treatment periods. Subjects will be maintained on DB dose and transition to the LTSE on the same dose.

After the Day 1 Visit, subsequent clinic visits during the double-blind treatment period will occur at approximately Weeks 4 (±5 days), 8 (±5 days), and 12 (±2 weeks) (efficacy, safety, tolerability, and PK will be assessed) and then every 12 weeks for the assessment of efficacy, safety, and tolerability. PK will not be assessed after Week 12. Subjects will also be contacted by telephone at Weeks 2 and 6 (±5 days) to assess for occurrence of any adverse events (AEs), changes to concomitant medications and/or new medications that have been initiated, and medical/surgical procedures, and to verify that the subject is dosing as directed.

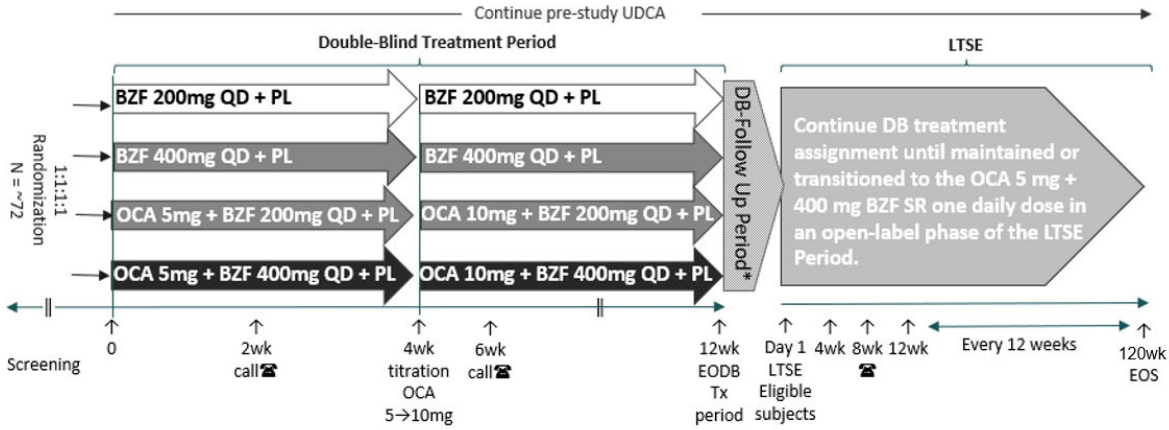
Subjects will transition to the LTSE period when the last subject completes Week 12 of the DB Period. Subjects will continue the original treatment assignment allocated during the DB Period until maintained or transitioned to OCA 5 mg + 400 mg BZF SR once daily in an open-label phase of the LTSE period. All site staff will maintain study blind until the study is moved to the open-label phase of the LTSE Period.

An independent Data Monitoring Committee (DMC) will review safety data and other relevant information from this study for their recommendation. Ad hoc meetings will be convened, as appropriate. The role and scope of the DMC will be further described in their separate charter.

The database locks for the analysis at the end of DB period will be a final type of database lock.

3.2. Study Design Diagram

Figure 1: Study Design for the Double-Blind/Follow-Up and LTSE Treatment Periods (Study 747-213)



BZF=bezafibrate; DB=double-blind; EODB=end of DB; EOS=end of study/end of LTSE Period; FDC=fixed dose combination; LTSE=long-term safety extension; OCA=obeticholic acid; LTSE=long-term safety extension); QD=once daily; UDCA=ursodeoxycholic acid. PL=Placebo=either OCA or BZF tablets.

Note: *DB Follow-Up Period. Subjects will move into the DB Follow-Up Period and remain in this period until the last subject finishes the DB Period.

Note: LTSE Period: Subjects will transition to the LTSE Period when the last subject completes Week 12 of the DB Period. Subjects will continue the original treatment assignment allocated during the DB Period until maintained or transitioned to the OCA 5 mg + 400 mg BZF SR once daily dose in an open-label phase of the LTSE Period.

Note: The Screening period is 2 weeks to a maximum of 8 weeks.

Note: Subjects taking UDCA at the time of enrollment will remain on their stable dose of UDCA during the study.

Note: All patients will transition to LTSE Day 1 at the same time from either the DB Period (Week 12; last subject completing the DB period) or the DB Follow-Up Period (remainder of all subjects).

3.3. Randomization and Blinding

This study will be conducted in a double-blind, randomized manner on subjects enrolled under current protocol approved at the site. Enrolled subjects will be randomized in a 1:1:1:1 ratio to receive one of the following treatment groups:

Treatment A: BZF 200 mg IR tablet QD	OCA 5 mg placebo tablet (Weeks 0 to 4) OCA 10 mg placebo tablet (Week 4 to end of DB treatment) BZF 200 mg IR tablet BZF 400 mg SR placebo tablet
Treatment B: BZF 400 mg SR tablet QD	OCA 5 mg placebo tablet (Weeks 0 to 4) OCA 10 mg placebo tablet (Week 4 to end of DB treatment) BZF 400 mg SR tablet BZF 200 mg IR placebo tablet
Treatment C: OCA 5 mg→10 mg QD + BZF 200 mg IR QD	OCA 5 mg tablet (Weeks 0 to 4) OCA 10 mg tablet (Week 4 to end of DB treatment) BZF 200 mg IR tablet BZF 400 mg SR placebo tablet
Treatment D: OCA 5 mg→10 mg QD + BZF 400 mg SR QD	OCA 5 mg tablet (Weeks 0 to 4) OCA 10 mg tablet (Week 4 to end of DB treatment) BZF 200 mg IR placebo tablet BZF 400 mg SR tablet

BZF=bezafibrate; DB=double-blind; IR=immediate release; OCA=obeticholic acid; QD=once daily; SR=sustained release

All subjects will be assigned investigational product according to their treatment group from Day 1 using an interactive web response system (IWRS).

Randomization will be stratified at Baseline by the following:

- Total bilirubin levels (≤0.7x or >0.7x ULN) but <2x ULN
- ALP ≥1.5x ULN but ≤1.67x ULN (not more than 25% of subjects enrolled)
- ALP >1.67x ULN (remaining subjects enrolled)

The unblinded data access during DMC and interim analysis is described in DMC Charter and Data Access Plan.

3.4. Sample Size Determination

[redacted] Assuming the change from baseline in ALP with BZF treatment is -146 U/L, a sample size of 18 subjects per arm will provide 80% power to detect a treatment difference of 54 between the BZF and BZF+OCA arms at two-sided alpha of 0.05 with the pooled standard deviation of 58.

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 be the population used for the efficacy endpoints. ITT population will be used in the primary analysis of primary efficacy endpoint. Treatment assignment will be based on the randomized treatment.

4.2. Modified Intent-to-Treat Population

The modified Intent-to-Treat (mITT) Population will include all randomized subjects who have baseline and at least one post baseline ALP assessment and will be the population used for the sensitivity analyses of primary efficacy endpoint. Treatment assignment will be based on the randomized treatment.

4.3. The Safety Population

The Safety Population will include all randomized subjects who receive at least 1 dose of OCA and/or BZF. The Safety Population will be the population used for safety analyses. Treatment assignment will be based on the actual treatment received.

4.4. PK Population

The PK Population will include all subjects who receive OCA and/or BZF and have enough quantifiable PK samples without any major protocol deviations that could potentially affect plasma exposure. The PK Population will be used for PK and PD analyses.

4.5. The LTSE Population

The LTSE Population will include all subjects who receive at least 1 dose of BZF and/or OCA during the LTSE period after the OLE transition. The LTSE Population will be used for LTSE analyses.

5. GENERAL CONSIDERATION

Individual subject data obtained from electronic case report forms (eCRFs), central and local laboratories, external sources, and any derived data will be presented in data listings by subject. All data listings that contain an evaluation date will contain a relative study day. Pre-treatment and on -treatment study days are numbered relative to the day of the first dose of investigational product which is designated as Day 1. The preceding day is Day -1, the day before that is Day 2, etc. The last day of investigational product is designated- with an "L" (e.g., Day 14L). Post-treatment study days are numbered relative to the last dose and are designated as Day 1P, Day 2P, etc.

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

Tabulations will be produced for appropriate demographic, baseline, efficacy, and safety parameters.

For categorical variables, summary tabulations of the number and percentage of subjects within each category (with a category for missing data) of the parameter will be presented. Percentage calculations will be based on non-missing data, unless otherwise specified. Percentages are rounded to 1 decimal place, unless otherwise specified.

For continuous variables, the number of subjects, mean, standard deviation (SD), standard error of the mean (SEM), median, minimum, and maximum values will be presented.

All statistical tests comparing groups will be conducted at the 2-sided, 0.05 level of significance. Summary statistics for each treatment group will be presented, as well as 95% CIs comparing groups will be provided.

The subject enrolled under the version 2 protocol will be presented in the listings but will not be included in any tables and figures.

The DB period is defined as the period from the first DB dose date to the last date before the OLE transition.

The LTSE period is defined as the period from the OLE transition to the last visit date in the study.

5.1. Baseline Definitions

The Baseline value for statistical analyses in ITT, mITT and safety populations is defined as the last evaluations prior to the first administration of investigational product, unless otherwise specified.

The baseline value for statistical analyses in LTSE population is defined as the last assessment on or before the OLE transition date.

5.2. Partial Dates

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

- Diagnostic Date (e.g., PBC diagnostic date)
  - For missing day only – Day will be imputed as the first day of the month (i.e., 1).
  - For missing day and month – Day and month will be imputed as the first day of the year (i.e., 1 January).
- Start Dates (e.g., 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 (i.e., 1) with the following exception: if the partial date falls in the same month and year as the date being used in the calculation (e.g., 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 (i.e., 1 January) with the following exception: if the partial date falls in the same year as the date being used in the calculation (e.g., 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.
- Stop Dates (e.g., AE resolution date or stop date of medication)
  - For missing stop day only – Day will be imputed as the last day of the month (i.e., 28, 29, 30, or 31).
  - For missing stop day and month – Day and month will be imputed as the last day of the year (i.e., 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.

All data recorded on the eCRF will be included in data listings that will accompany the CSR.

5.3. Data Conventions

The precision of original measurements will be maintained in summaries, when possible.

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

Percentages based on frequency counts will be based on available data, and denominators will generally exclude missing values, unless otherwise stated. For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness 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%."

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

Standard 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

5.4. Calculations

Variables requiring calculation will be derived using the following formulas:

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

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

  duration in days=date2 – date1, where date1 <first dose date

- Months – A duration expressed in months is calculated as the number of days divided by 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

- Age of diagnosis – Age of diagnosis is calculated as the number of years from the imputed year of birth (*DOB*) to the year of diagnosis. The imputed year of birth is derived as year of inform consent – age. The following formula is used:

  age (years)=year of diagnosis – imputed year of DOB + 1

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

  height (cm)=height (in) × 2.54

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

  weight (kg)=weight (lb) / 2.2046

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

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

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

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

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

  Change=post Baseline value – Baseline value

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

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





5.5. Analysis Windows

Analysis windows will not be defined in this study. Only the scheduled visits will be included in the analyses. All visits, including unscheduled visits, will be presented in the listings.

6. SUBJECT DISPOSITION

Subject disposition will be tabulated by treatment group and overall and will include the following number and proportion of subjects for Double-Blind Period:

- Randomized
- Treated (Safety)
- mITT population
- Per protocol population
- PK population
- Completed study
- Study discontinuation and reasons
- Investigational product discontinuation and reasons

Subject disposition will also be tabulated by Double-Blind treatment group and overall and will include the following number and proportion of subjects for LTSE Period:

- Completed study
- Investigational product discontinuation and reasons

Disposition summaries of both end of study and end of treatment will be presented for each analysis population and will be tabulated by treatment group and overall. Proportions will be out of the number of randomized subjects.

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

6.1. Protocol Deviations

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

Major protocol deviations that could potentially affect the efficacy or safety conclusions of the study will be identified prior to database lock and unblinding of individual subject treatment information. Major protocol deviations will be summarized by deviation category and treatment group using ITT population.

Protocol deviations will be presented in a by-subject data listing.

7. DEMOGRAPHICS AND BASELINE DISEASE CHARACTERISTICS

7.1. Demographics and Baseline Characteristics

Demographic variables will include the following:

- Age at informed consent (Continuous and Categorical (<65 and ≥65))
- Sex (Male and Female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White and Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported, Unknown)
- Geographic Region

Other Baseline characteristics will include the following:

- Weight (kg)
- Height (cm)
- BMI (kg/m^2)

Key liver function test results as a continuous variable and categorized as follows:

- ALP (U/L)
- Total Bilirubin (μmol/L)
- Albumin (g/L)
- Use of UDCA (Yes/No)

- Total daily dose of UDCA (mg) as determined using the last dosage reported prior to the first dose

Demographics and Baseline characteristics will be summarized and presented by treatment group and overall for ITT, Safety and PK populations. 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, SD, SEM, median, 25th and 75th quartiles, minimum, and maximum values. No inferential statistical comparisons will be performed.

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

7.2. Baseline PBC Disease Characteristics

Baseline PBC disease characteristics will be summarized using data collected from the PBC Disease History and other eCRFs. Variables of the PBC Disease History include the following:

- Age at PBC diagnosis
- Duration of PBC in years at time of informed consent
- History of PBC related pruritus (Yes/No)
- Severity of most recent pruritus event
  - Ongoing at screening
- History of PBC-related fatigue (Yes/No)
  - Ongoing at screening
- Overall severity of PBC-related fatigue

Baseline PBC disease characteristics will be summarized and presented by treatment group and overall for ITT, Safety and PK populations.

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

8. MEDICAL HISTORY

Verbatim terms on eCRFs will be mapped to preferred terms and system organ classes using the Medical Dictionary for Regulatory Activities (MedDRA).

Medical history will be summarized by system organ class, preferred term, and treatment group using the Safety population. Summaries will be ordered by descending order of incidence of system organ class and preferred term within each system organ class.

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

9. EFFICACY ANALYSES

The efficacy endpoints of the study are defined below. Additionally, the planned analyses for these endpoints are described. All efficacy endpoints will be presented in by-subject data listings.

Subjects who discontinue investigational product during the double-blind treatment period are expected to continue until the end of double-blind treatment period of this study.

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

During the conduct of the study, the DMC may recommend refining the missing data strategy to better address the observed pattern of missing data, based on blinded monitoring of the data.

All efficacy analyses will be summarized using descriptive statistics at Baseline and at each scheduled postbaseline visit. The change from baseline and percent change from baseline will also be summarized. Descriptive statistics, including change from baseline, percent change from baseline, and estimates of least-square (LS) means, SEs, and 95% CIs, will be presented by treatment group. Estimates of the mean difference between treatment groups, the SE of the difference, and 2-sided 95% CI of the difference will also be presented.

For quantitative endpoints evaluated using a mixed-effect repeated measures model (MMRM), no imputation will be performed on the missing data.

For binary response endpoints analyzed by Cochran Mantel Haenszel (CMH) test, two separate analyses will be done. One analysis will impute the missing data as non-responder, and the other analysis will exclude the missing subjects. In the case that more than 20% cells with frequencies <5, a fisher exact test will be used to compare the treatment groups.

In an efficacy analysis in which subjects are classified as either a responder or a non-responder (binary outcome) based on dichotomizing a continuous variable, any subject who does not provide an assessment at the specified time point for the defining of response will be considered to be 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.

For supportive analyses using only “observed cases”, subjects who do not provide an assessment at the specified timepoint for the defining of response will not be included. That is, for the percentage of responders, the subject will not be included in the numerator or the denominator.

9.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the change in ALP from baseline to Week 12 in the DB Period. The primary efficacy analysis will compare subjects treated with Treatment C versus Treatment A and Treatment D versus Treatment B with respect to the primary efficacy endpoint using ITT population in DB Period. For the primary efficacy analysis, the subjects with baseline and at least one post-baseline ALP will be included in the analysis.

The intercurrent event in this study includes the use of prohibited medication and treatment discontinuation. For the primary analysis on the primary efficacy endpoint, data collected after the intercurrent events, such as the use of prohibited medications, and data collected after treatment discontinuation will be included.

Analyses of change in ALP will be evaluated using a MMRM to evaluate the effect of treatment groups over time. The dependent variable will be the change from Baseline. The model will

include the terms of treatment group, visit, treatment by visit interaction, randomization stratification variables derived from lab parameters as factors and baseline value as covariate.

An unstructured covariance model will be used. If the estimation procedure fails to converge, the following structures will be tested in the order: Toeplitz, first-order autoregressive (AR [1]) and compound symmetry (CS).

If the model still fails to converge, the stratification factor of ALP (≥1.5x ULN but ≤1.67x ULN or >1.67x ULN) will be dropped from the model. The model will be tested using unstructured covariance model with reduced parameter in the model. If the estimation procedure fails to converge, the simpler structures will be tested again in the above order.

If the model still fails to converge, the stratification factor of total bilirubin levels (≤0.7x or >0.7x ULN) but <2x U LN) will be dropped from the model too. The model will be tested using unstructured covariance model with none of the stratification factors. If the estimation procedure fails to converge, the simpler structures will be tested again in the above order.

The Kenward and Roger method will be used to calculate the denominator degrees of freedom for the test of fixed effects.

Estimates of LS means, SEs, and 95% CIs will be presented by treatment group. Estimates of the mean difference between treatment groups, the SE of the difference, and 95% CI of the difference will be presented. Comparison of treatment groups will be based on their mean estimates, the associated 95% CIs and p-value.

The results provided by MMRM are based on the assumption that the residuals from the fitted model follow the normal distribution. This assumption will be tested for the primary efficacy variable, using the Shapiro-Wilk test at a 2-sided alpha of 0.05 significance level. In the case of a significant result, normal plots will be used to assess the degree of deviation from normality. If a substantial deviation from normality is detected, a supplemental analysis of covariance (ANCOVA) on ranks stratified by the same stratification variables will be performed on week 12, computing p-values for comparing treatment groups.

In addition, a supportive analysis using ITT population will be performed on the primary efficacy endpoint, where data collected after the intercurrent events will not be included in the analysis. This provides the evaluation of efficacy effect according to treatment adherence.

The primary efficacy analyses will also be conducted using the mITT Population as sensitivity analyses.

9.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

- ALP response rates of ≥10%, ≥20%, ≥30% and ≥40% reduction from baseline and normalization rates at Week 12
- Normalization rates at Week 12 of GGT, ALT, AST, ALP, total and conjugated bilirubin, and a lipid panel
- Change from baseline to Week 12 in GGT, ALT, AST, and total and conjugated bilirubin, and a lipid panel

- Change from baseline to Week 12 in 7α-hydroxy-4-cholesten-3-one (C4) and bile acids

For the secondary efficacy endpoints, data collected after the intercurrent events, such as the use of prohibited medications, and data collected after treatment discontinuation will be included.

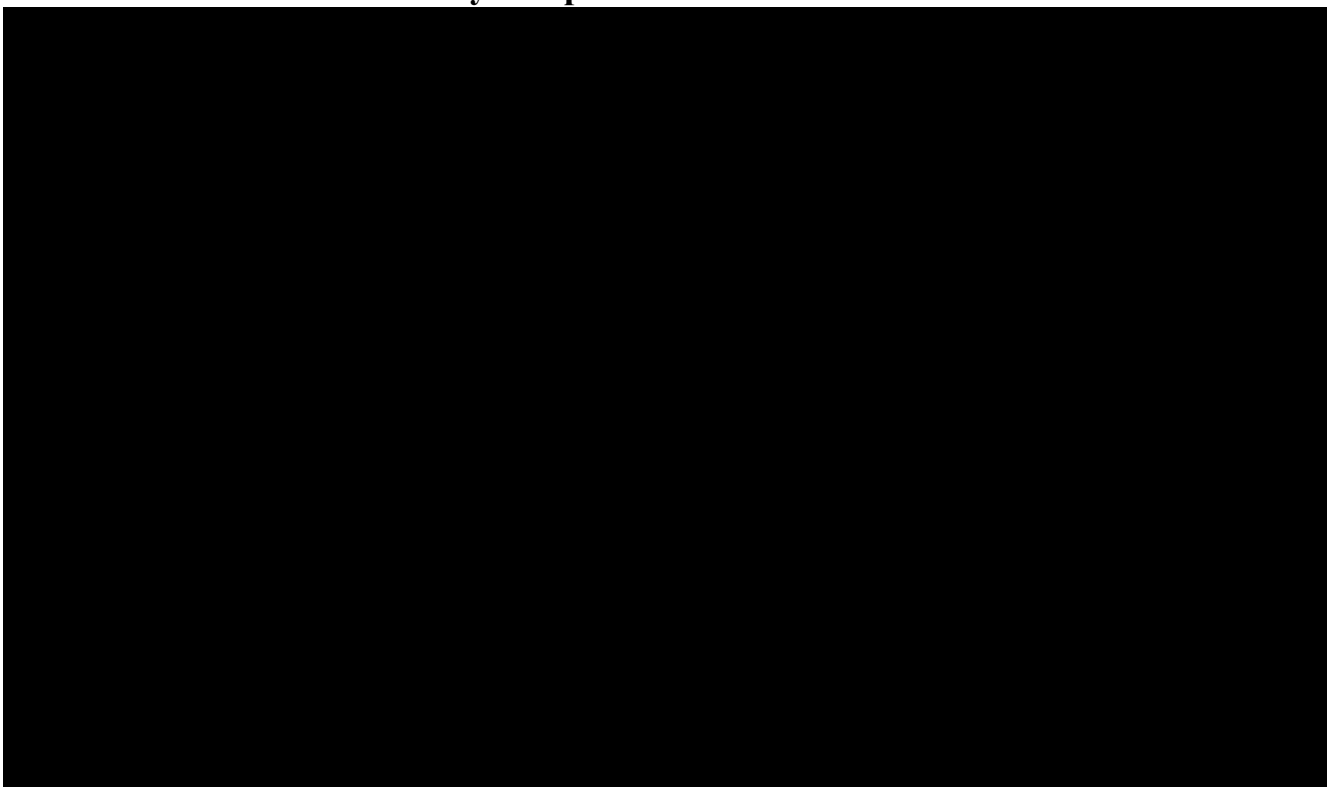
The secondary efficacy analysis will compare subjects treated with Treatment C versus Treatment A and Treatment D versus Treatment B with respect to the secondary efficacy endpoints using the ITT population. The secondary efficacy analyses will not be analyzed in the mITT Population.

Analyses for ALP response rates of ≥10%, ≥20%, ≥30%, and ≥40% reduction and normalization rates will compare Treatment C versus Treatment A and Treatment D versus Treatment B using a CMH test stratified by the randomization stratification factors derived from lab parameters.

Longitudinal endpoints will be analyzed using the same MMRM model as specified in Section 9.1. The dependent variable will be the change from Baseline. If the normality assumption is not satisfied, a supplemental analysis of ANCOVA on ranks described in Section 9.1 will be performed.

Summary statistics of the laboratory values at each visit will be provided by treatment group. The results, change from Baseline, and percentage change from Baseline values, as well as estimates of LS means, SE, and 95% CIs, will be presented by treatment group. Estimates of the mean difference between treatment groups, the SE of the difference, and 95% CI of the difference will also be presented.

9.3. Additional Efficacy Endpoints



[redacted]

9.3.1. Quantitative Endpoints

[redacted]

9.3.2. Prognostic Endpoints

9.3.2.1. Change from Baseline in Mayo Risk Score

The Mayo Risk Score (MRS) is recorded on the eCRF and calculated as: 0.871 ln(bilirubin; mg/dL) - 2.53 ln (albumin; g/dL) + 0.039 (age; yr) + 2.38 ln (PT; sec) + 0.859 (edema; 0 for none, 0.5 if resolved with diuretics, 1 if unresolved with diuretics). Higher scores on MRS indicate worse disease severity.

MRS during treatment will be compared to Baseline values to determine if there was any increase in MRS scores. This information will be reported by visit. Change from Baseline in MRS scores will be summarized and analyzed using the same MMRM model as specified in Section 9.1.

9.3.2.2. Child-Pugh

The Child-Pugh (CP) score is calculated by adding the scores of the 5 factors and can range from 5-15. Child-Pugh class is either A (a score of 5-6), B (7-9), or C (10 or above). Each of the factors is outlined in Table 1.

Table 1: Child-Pugh Score

Factor	Units	Points		
		1	2	3
Serum bilirubin	µmol/L	<34	34-50	>50
	mg/dL	<2.0	2.0-3.0	>3.0
Serum albumin	g/L	>35	35-28	<28
	g/dL	>3.5	3.5-2.8	<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 encephalopathy		No	Grade 1 or 2	Grade 3 or 4

INR=international normalized ratio

The CP score will be summarized in the same manner as described above for the [redacted]

Child-Pugh class will be presented in shift tables showing the shift from baseline to each scheduled post-baseline visit by the count and percentage of subjects within each shift classification.

Individual factors will be presented in a listing.

9.4. Analyses of LTSE Efficacy Endpoints

All efficacy endpoints in LTSE period will be summarized by visits. No statistical inference will be done on those endpoints.

9.5. Interim Analyses and Data Monitoring Committee

9.5.1. Interim Analyses

The interim analysis will be conducted by a Sponsor sub-team that will not be part of the study level activities after the interim analysis. The data access plan (DAP) will specify the make-up of the Sponsor sub-team and the appropriate data firewalls to ensure data integrity. A small group at Intercept Pharmaceuticals will also be unblinded. The details will be specified in the Data Unblinding Plan.

The interim analysis will be carried out when at least 40 subjects complete the double-blind treatment period of the study. In addition to the routine safety review, the DMC will also review the interim analysis.

The Sponsor will have the option to perform an additional unblinded rolling analysis of efficacy, PK, and safety data throughout the DB Period. Details of the unblinding and DAP, including Sponsor staff that will be unblinded, will be provided. All Principal Investigators, site staff, and subjects will remain blinded throughout the study.



Analyses of those response rates will compare Treatment C versus Treatment A and Treatment D versus Treatment B using a

The changes from baseline to Week 12 in GGT, ALT, AST, total bilirubin and conjugated bilirubin in the ITT population will be analyzed using the same MMRM model as ALP.

The analyses of drug exposure and safety assessments, including the numbers of subjects with pruritus events will also be performed.

Those endpoints will be summarized using descriptive statistics at Baseline and at each scheduled post-baseline visit.

9.5.2. Data Monitoring Committee

An independent DMC will review safety and efficacy data from this study for their recommendation. Ad hoc meetings will be convened, as appropriate. The role and scope of the DMC will be further described in a separate DMC charter.

9.6. Multiple Comparisons/Multiplicity

No adjustment on type I error will be performed in this study. All p-values provided in the tables should be considered nominal.

9.7. Subgroup Analyses of Efficacy

No subgroup analysis of efficacy is planned in this study.

10. SAFETY ANALYSES

Safety analyses will be performed using the Safety population including by treatment groups and pooled across all active dose levels. No inferential comparison of safety endpoints will be performed. The analyses for safety data within the first 12 weeks of DB period, in the whole DB treatment period and LTSE periods will be analyzed separately.

10.1. Exposure to Study Treatment

The extent of exposure will be summarized using descriptive statistics.

Duration 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

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

- Duration of temporary discontinuation of investigational product=(Date of restart of investigational product – Date of temporary discontinuation of investigational product) + 1

The total duration of temporary investigational product discontinuation is the sum duration of temporary discontinuation of investigational product over each incidence of discontinuation.

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

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

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

where

- number of tablets consumed during study=number of tablets dispensed-number of tablets returned

Subject compliance with investigational product will be summarized by treatment group using descriptive statistics. All exposure data will be presented in a by subject data listing.

10.2. Adverse Events

AEs are defined as any untoward medical occurrence associated with the use of the investigational product in humans, whether or not considered related to the investigational product. An AE (also referred to as an adverse experience) can be any unfavorable and unintended sign (e.g., an abnormal laboratory finding), symptom, or disease temporally

associated with any use of the investigational product, without any judgment about causality and irrespective of route of administration, formulation, or dose, including an overdose.

AEs will be mapped to preferred terms (PTs) and system organ classes (SOCs) using MedDRA dictionary version 22.0. 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.

10.2.1. Treatment-emergent Adverse Events

An AE is defined as a treatment-emergent adverse event (TEAE) within the first 12 weeks of the Double-Blind period if it meets one or more of the following criteria:

- An AE starting during the period between the first study drug dose date and the week 12 visit date in Double-Blind period (inclusive).
- An AE occurring prior to the first study drug dose that worsens (increase in grade) during the period between the first study drug dose date and the week 12 visit date in Double-Blind period (inclusive).
- An AE with a missing start date but with an end date between the first study drug dose date and the week 12 visit date in Double-Blind period (inclusive).
- Any AE started either later than the week 12 visit date in Double-Blind period or after the first dose of LTSE period is not counted as a TEAE in the first 12 weeks of Double-Blind period.

An AE is defined as a treatment-emergent adverse event (TEAE) in the whole DB period if it meets one or more of the following criteria:

- An AE starting on or after the first study drug dose of Double-Blind period.
- An AE occurring prior to the first study drug dose that worsens (increase in grade) after the first study drug dose during Double-Blind period.
- An AE with a missing start date but with an end date after the first study drug dose date.
- Any AE started either later than the 30 days of the last dose of study drug in Double-Blind period or after the first dose of LTSE period is not counted as a TEAE in the Double-Blind period.

An AE is defined as a TEAE in the LTSE period after OLE transition if it meets one or more of the following criteria:

- An AE starting on or after the first day of OLE transition during LTSE period and within the 30 days after the last dose of study drug of LTSE period.
- An AE occurring prior to the first study drug dose of OLE transition during LTSE period that worsens (increase in grade) after the first study drug dose of OLE transition during LTSE period.

10.2.2. 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 or incapacity
- Results in a congenital abnormality or birth defect
- 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 emergency 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.

Events **not** considered to be SAEs 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
- Respite care or observation when there is no AE associated with the hospitalization

10.2.3. Adverse Events of Special Interest

The Adverse Events of Special Interest (AESIs) include pruritus, hepatic injury, myalgia and renal impairment.

10.2.4. TEAE Analyses

All AE summaries will be restricted to TEAEs. Each AE summary will be displayed by treatment group for DB and LTSE periods separately. Summaries that are displayed by SOC and PT will be ordered by descending order of incidence of system organ class and preferred term within each system organ class. Summaries of the following types will be presented:

- Overall summary of TEAEs
- Subject incidence of TEAEs and the total number of entries by MedDRA SOC and PT.
- Subject incidence of TEAEs by MedDRA SOC, PT, and highest severity. 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.
- Subject incidence of TEAEs by MedDRA SOC, PT, and closest relationship to investigational product (Related/Not Related). Related AEs are those with relationships reported as “Definite,” “Probable,” or “Possible,” and unrelated AEs are those with relationships 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 a missing relationship will be considered related for this summary.

- Subject incidence of serious TEAEs and the total number of entries by MedDRA SOC and PT.
- Subject incidence of TEAEs leading to investigational product withdrawal or study discontinuation by MedDRA SOC and PT. 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.

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

- All adverse events
- Serious adverse events (a subset of the AEs where serious is marked as "Yes").
- Severe adverse events (a subset of AEs where severity is marked as "Severe").
- Related adverse events (a subset of the AEs where relationship marked as "Definite," "Probable," or "Possible").
- Adverse events leading to Study Drug Withdrawal or Study Discontinuation (a subset of the AEs where Action Taken with Study Treatment is checked as "Drug Withdrawn" or where "Subject Discontinued from Study" is checked).
- Adverse events leading to death.

10.2.4.1. Pruritus

Pruritus is both the most common symptom of PBC and was the most frequently reported AE in the Phase 2 PBC studies for OCA. Accordingly, pruritus is considered an AE of special interest.

Treatment-emergent pruritus, defined as any preferred term including "Prur," will be summarized separately by the MedDRA SOC, treatment group, and PT as a subset of all TEAEs. Summaries will include a summary of the crude incidence of Pruritus TEAEs and SAEs.

An analysis of treatment-emergent pruritus event days per patient years on study will be performed for DB and LTSE periods separately. 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 order to explore the relationship between pruritus and the investigational product, the following time-to-event analyses will be performed:

- Time to first onset of treatment-emergent pruritus
  - The time to the start of the first episode will be calculated by date of onset of first episode – date of first dose of investigational product + 1
  - Subjects who never reported an AE of pruritus will be censored at the date of last contact

- Time to onset of the first moderate or severe treatment-emergent pruritus
  - The time to the start of the first moderate or severe pruritus will be calculated by date of onset of the first moderate or severe pruritus – date of first dose of investigational product +1
  - Subjects who never reported a moderate or severe AE of pruritus will be censored at the date of completion or discontinuation
- Time to onset of the first severe treatment-emergent pruritus
  - The time to the start of the first severe pruritus will be calculated by date of onset of the first severe pruritus – date of first dose of investigational product +1
  - Subjects who never reported a severe AE of pruritus will be censored at the date of completion or discontinuation

The analysis of time to first onset of treatment-emergent pruritus AE, time to onset of the first moderate or severe treatment-emergent pruritus and onset of the first severe treatment-emergent pruritus will include the number of subjects with pruritus (first onset, first moderate or severe, first severe), the number of subjects without pruritus (censored), and the minimum and maximum times in days. Kaplan-Meier (KM) estimates will be calculated by treatment group. The quartiles, including the median time-to-event and their respective 2-sided 95% CIs, will be presented. KM estimates will be plotted as a "survival curve" for each treatment group.

10.2.4.2. Hepatic Injury

Hepatic injury is defined as events included in the Hepatic Disorder Standardized MedDRA query (SMQ), excluding the following sub-SMQs: alcohol related, congenital, familial, neonatal and genetic disorders of the liver; liver infections; and pregnancy-related hepatic injury. Summaries will include a summary of the crude incidence of hepatic injury TEAEs and SAEs will be reported.

In order to explore the relationship between hepatic injuries and investigational products, the following time-to event analyses will be performed:

- Time to first onset of treatment emergent hepatic injuries
  - The time to the start of the first onset will be calculated by date of onset of first onset – date of first dose of investigational product + 1
  - Subjects who never reported a hepatic injury will be censored at the date of last contact

KM estimates will be calculated by treatment group. The quartiles, including the median time to event and their respective 2-sided 95% CIs, will be presented. Figures for time to first onset will be provided, with the number at risk identified.

10.2.4.3. Myalgia

Myalgia is defined as events with the preferred term of Myalgia. Summaries will include a summary of the crude incidence of myalgia. TEAEs and SAEs will be reported.

In order to explore the relationship between hepatic injuries and investigational products, the following time-to event analyses will be performed:

- Time to first onset of treatment emergent hepatic injuries
  - The time to the start of the first onset will be calculated by date of onset of first onset – date of first dose of investigational product + 1
  - Subjects who never reported a hepatic injury will be censored at the date of last contact

KM estimates will be calculated by treatment group. The quartiles, including the median time-to-event and their respective 2-sided 95% CIs, will be presented. Figures for time to first onset will be provided, with the number at risk identified.

10.2.4.4. Renal Impairment

Renal impairment is defined as events included in the following broad SMQs: acute renal failure, chronic kidney disease, proteinuria, renovascular disorders, or tubulointerstitial diseases.

Summaries will include a summary of the crude incidence of myalgia TEAEs and SAEs will be reported.

To explore the relationship between renal impairment and investigational products, the following time-to event analyses will be performed:

- Time to first onset of treatment emergent renal impairment
  - The time to the start of the first onset will be calculated by date of onset of first onset – date of first dose of investigational product + 1.
  - Subjects who never reported a renal impairment will be censored at the date of last contact.

KM estimates will be calculated by treatment groups. The quartiles, including the median time-to-event and their respective 2-sided 95% CIs, will be presented. Figures for time to first onset will be provided, with the number at risk identified.

10.3. Clinical Laboratory Evaluations

Laboratory parameters will be summarized in standard international system of units by treatment group using descriptive statistics at Baseline and at each on-study evaluation. Re-test or unscheduled results will not be summarized for DB and LTSE periods separately but will be included in the data listings. The change from Baseline will also be summarized for the first 12 weeks of DB period, the whole DB period and LTSE period separately. Baseline is defined as the mean of all available evaluations prior to the first administration of investigational product.

The normal ranges of the selected lab parameters are defined below. The abnormal indicators of the selected lab parameters will be derived based on these normal ranges (inclusively):

- ALP 44-122 U/L
- ALT 0-45 U/L
- AST 0-45 IU/L
- GGT 0-65 IU/L (male) 0-60 IU/L (female)
- Total Bilirubin 1.71-12.3 umol/L
- Direct Bilirubin 0-0.4 mg/dL
- Cholesterol 0-169 mg/dL (age <=19), 0-199 mg/dL (age >19)
- LDL 0-109 mg/dL (age <=19), 0-99 mg/dL (age >19)
- HDL >39 mg/dL
- Creatinine 0.57-1.27 mg/dL
- Creatine 0.0-0.7 mg/dL (male) 0.1-1.0 mg/dL (female)
- CK 41-439 U/L (male, age 18-50 years old), 41-331 U/L (male, age 51-80 years old),
- 32-182 (female, age 18-80 years old)

For the other lab parameters, the normal ranges and abnormal indicators are based on the values in the data from the central and local labs.

In addition, shift tables (i.e., low-normal-high at Baseline versus low-normal-high at post-Baseline visit in a 3-by-3 contingency table) from Baseline to worst value, last value, and at each scheduled post Baseline visit will be provided for hematology and serum chemistry by treatment group to assess changes in laboratory values from Baseline to each on-study evaluation. The subjects meeting Hy's law criteria (defined as AST or ALT ≥3 xULN together with total bilirubin ≥2 xULN at any timepoint) will be summarized by treatment and visit for the first 12 weeks of DB period, the whole DB period and LTSE period separately.

Urine chemistry and urinalysis results will not be summarized but will be provided in a data listing.

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

10.4. Vital Signs

The results and change from Baseline to each on-study evaluation visit will be summarized for weight, oral temperature, sitting heart rate, respiratory rate and sitting blood pressure for the first 12 weeks of DB period, the whole DB period and LTSE period separately.

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

10.5. Electrocardiograms

Overall interpretation results for electrocardiograms (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. The results of ECG will be summarized by visit for the first 12 weeks of DB period, the whole DB period and LTSE period separately.

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

10.6. Physical Exam

The results of physical exam will be summarized by visit for the first 12 weeks of DB period, the whole DB period and LTSE period separately.

All physical exam data will be presented in by-subject data listings.

10.7. Prior and Concomitant Medications

Prior medications are those medications with start and stop prior to the initial dose of investigational product. Concomitant medications are those medications started prior and continued after the initial dose of investigational product were. If it cannot be determined whether the medication was a 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 concomitant medication.

Prior medications will be listed only. Concomitant medications will be summarized for each treatment group by World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) class level 2, WHO ATC class level 4 and preferred name for the first 12 weeks of DB period, the whole DB period and LTSE period separately using the Safety and LTSE Populations. These summaries will present the number and percentage of subjects using each medication. Subjects may have more than one medication per ATC class 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 class and preferred name within each ATC class.



[redacted]

10.9. Subgroup Analyses of Safety

No subgroup analysis of safety is planned in this study.

[redacted]

13. DEVIATIONS FROM PROTOCOL

There have been no deviations between the protocol-defined statistical analyses and those presented in this statistical analysis plan.

14. REFERENCES

Corpechot C, Chazouilleres O, Rousseau A, et al. A Placebo-Controlled Trial of Bezafibrate in Primary Biliary Cholangitis. The New England journal of medicine. 2018 Jun 7;378(23):2171-81.

APPENDIX A. DATA PRESENTATION CONVENTIONS:

Unless otherwise specified, the baseline value for analyses of all parameters is defined as the last non-missing value prior to the first administration of investigational product.

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.

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.

All statistical analyses will be performed using SAS statistical software Version 9.4 or higher unless otherwise noted.

APPENDIX B. SUMMARY OF CHANGES:

Rationale and Summary of Changes

The revisions to SAP Version 1.0 summarized in this appendix include the following:

Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
10	The analyses for DB and LTSE periods will be analyzed separately.	The analyses for safety data within the first 12 weeks of DB period, in the whole DB treatment period and LTSE periods will be analyzed separately.	A new set of safety tables are added to provide the safety summary for the first 12 weeks in DB period.
10.2.1	NA	An AE is defined as a treatment-emergent adverse event (TEAE) within the first 12 weeks of Double-Blind period if it meets one or more of the following criteria: • An AE starting during the period between the first study drug dose date and the week 12 visit date in Double-Blind period (inclusive). • An AE occurring prior to the first study drug dose that worsens (increase in grade) during the period between the first study drug dose date and the week 12 visit date in Double-Blind period (inclusive). • An AE with a missing start date but with an end date between the first study drug dose date and the week 12 visit date in Double-Blind period (inclusive). • Any AE started either later than the week 12 visit date in Double-Blind period or after the first dose of LTSE period is not counted as a TEAE in the first 12 weeks of Double-Blind period.	The definition of TEAEs in the new set of safety tables for the first 12 weeks in DB period is added.

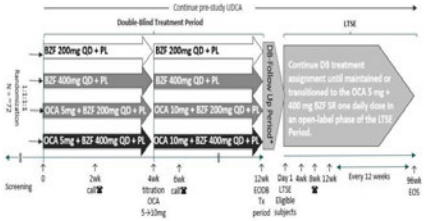
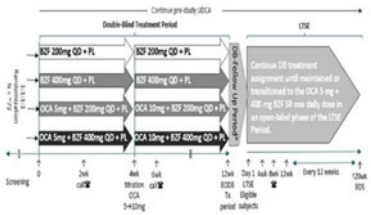
Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
10.3	NA	The normal ranges of the selected lab parameters are defined below. The abnormal indicators of the selected lab parameters will be derived based on these normal ranges: • ALP 44-122 U/L • ALT 0-45 U/L • AST 0-45 IU/L • GGT 0-65 IU/L (male) 0-60 IU/L (female) • Total Bilirubin 1.71-12.3 umol/L • Direct Bilirubin 0-0.4 mg/dL • Cholesterol 0-169 mg/dL (age <=19), 0-199 mg/dL (age >19) • LDL 0-109 mg/dL (age <=19), 0-99 mg/dL (age >19) • HDL >39 mg/dL • Creatinine 0.57-1.27 mg/dL • Creatine 0.0-0.7 mg/dL (male) 0.1-1.0 mg/dL (female) • CK 41-439 U/L (male, age 18-50 years old), 41-331 U/L (male, age 51-80 years old), 32-182 (female, age 18-80 years old) For the other lab parameters, the normal ranges and abnormal indicators are based on the values in the data from the central labs.	The normal ranges of a number of lab parameters are updated.

The revisions to SAP Version 2.0 summarized in this appendix include the following:

Section	Original Text (Version 2.0)	Revised Text (Version 3.0)	Key Change Reasons/ Justification for Change
5	Individual subject data obtained from electronic case report forms (eCRFs), central laboratories, external sources, and any derived data will be presented in data listings by subject.	Individual subject data obtained from electronic case report forms (eCRFs), central and local laboratories, external sources, and any derived data will be presented in data listings by subject.	The data from the local lab will also be included in the tables.

Section	Original Text (Version 2.0)	Revised Text (Version 3.0)	Key Change Reasons/ Justification for Change
[redacted]			
10.3	For the other lab parameters, the normal ranges and abnormal indicators are based on the values in the data from the central labs.	For the other lab parameters, the normal ranges and abnormal indicators are based on the values in the data from the central and local labs.	The data from the local lab will also be included in the tables.

The revisions to SAP Version 3.0 summarized in this appendix include the following:

Section	Original Text (Version 3.0)	Revised Text (Version 4.0)	Key Change Reasons/ Justification for Change
[redacted]			
3.1	Subjects will transition to the LTSE period when the last subject completes Week 12 of the DB Period. All site staff will maintain study blind and will continue the original treatment assignment allocated during the double-blind period.	Subjects will transition to the LTSE period when the last subject completes Week 12 of the DB Period. Subjects will continue the original treatment assignment allocated during the DB Period until maintained or transitioned to OCA 5 mg + 400 mg BZF SR once daily in an open-label phase of the LTSE Period . All site staff will maintain study blind until the study is moved to the open-label phase of the LTSE Period.	After Interim data review of the 747-213 DB period, the Sponsor has determined that the OCA 5 mg + 400 mg BZF once daily dose regimen is ideal. Therefore, this dose regimen has been selected for the LTSE period. Clarified language as all subjects will convert to the open-label phase of the LTSE Period per protocol v6
3.2			Extended LTSE Period to ensure patients could continue treatment until the Phase 2/3 LTSE protocol is initiated at study sites per protocol v7.

Section	Original Text (Version 3.0)	Revised Text (Version 4.0)	Key Change Reasons/ Justification for Change
3.2	Note: *DB Follow-Up Period. Subjects will move into the DB Follow-Up Period and remain in this period until the last subject finishes the DB Period. During the DB Follow-Up Period or LTSE, the dose of OCA and BZF may be optimized, and subjects will transition to OLE in the LTSE period.	Note: *DB Follow-Up Period. Subjects will move into the DB Follow-Up Period and remain in this period until the last subject finishes the DB Period. Note: LTSE Period: Subjects will transition to the LTSE Period when the last subject completes Week 12 of the DB Period. Subjects will continue the original treatment assignment allocated during the DB Period until maintained or transitioned to the OCA 5 mg + 400 mg BZF SR once daily dose in an open-label phase of the LTSE Period.	Extended LTSE Period to ensure patients could continue treatment until the Phase 2/3 LTSE protocol is initiated at study sites per protocol v6
4.3	**14.1 Per-Protocol Population** The Per-Protocol (PP) Population will include all subjects in ITT Population without any major protocol deviations. The PP Population will be used for the sensitivity analyses of primary efficacy endpoint. Treatment assignment will be based on the actual treatment received.	*Deletion*	Removed per protocol population due to small subject per protocol v6
7.1	Demographics and Baseline characteristics will be summarized and presented by treatment group and overall for ITT, Per-Protocol, Safety and PK populations.	Demographics and Baseline characteristics will be summarized and presented by treatment group and overall for ITT, Safety and PK populations.	Removed per protocol population due to small subject per protocol v6
7.2	Baseline PBC disease characteristics will be summarized and presented by treatment group and overall for ITT, Per-Protocol, Safety and PK populations.	Baseline PBC disease characteristics will be summarized and presented by treatment group and overall for ITT, Safety and PK populations.	Removed per protocol population due to small subject per protocol v6
9.1	The primary efficacy analyses will also be conducted using the mITT and Per-Protocol Population as sensitivity analyses.	The primary efficacy analyses will also be conducted using the mITT Population as sensitivity analyses.	Removed per protocol population due to small subject per protocol v6
9.2	Normalization rates at Week 12 of GGT, ALT, AST, total and conjugated bilirubin, and a lipid panel	Normalization rates at Week 12 of GGT, ALT, AST, ALP, total and conjugated bilirubin, and a lipid panel	Added the normalization rate of ALP in the analyses per protocol v6

Section	Original Text (Version 3.0)	Revised Text (Version 4.0)	Key Change Reasons/ Justification for Change
9.2	The secondary efficacy analyses will not be analyzed in the mITT and Per-Protocol Population.	The secondary efficacy analyses will not be analyzed in the mITT Population.	Removed per protocol population due to small subject per protocol v6
9.3	The additional efficacy analyses will not be analyzed in the mITT and Per-Protocol Population.	The additional efficacy analyses will not be analyzed in the mITT Population.	Removed per protocol population due to small subject per protocol v6
[redacted]			
10.2.3	The Adverse Events of Special Interest (AESIs) include pruritus, hepatic injury and myalgia.	The Adverse Events of Special Interest (AESIs) include pruritus, hepatic injury, myalgia and renal impairment.	Added renal impairment as adverse events of special interest (AESIs) to align with other FDC clinical studies per protocol v6

Section	Original Text (Version 3.0)	Revised Text (Version 4.0)	Key Change Reasons/ Justification for Change
10.2.4.4	*New Section*	Renal impairment is defined as events included in the following broad SMQs: acute renal failure, chronic kidney disease, proteinuria, renovascular disorders, or tubulointerstitial diseases. Summaries will include a summary of the crude incidence of myalgia TEAEs and SAEs will be reported. To explore the relationship between renal impairment and investigational products, the following time-to event analyses will be performed: • Time to first onset of treatment emergent renal impairment – The time to the start of the first onset will be calculated by date of onset of first onset – date of first dose of investigational product + 1. – Subjects who never reported a renal impairment will be censored at the date of last contact. KM estimates will be calculated by treatment groups. The quartiles, including the median time-to-event and their respective 2-sided 95% CIs, will be presented. Figures for time to first onset will be provided, with the number at risk identified.	Added renal impairment as adverse events of special interest (AESIs) to align with other FDC clinical studies per protocol v6