OBETICHOLIC ACID (OCA) AND BEZAFIBRATE (BZF)

Protocol 747-214

A Phase 2a, Double-Blind, Randomized, Active Controlled, Parallel Group Study Evaluating the Efficacy, Safety, and Tolerability of Bezafibrate Administered in Combination with Obeticholic Acid in Subjects with Primary Biliary Cholangitis

Statistical Analysis Plan

Author: [redacted]
[redacted]Biostatistics

Sponsor: Intercept Pharmaceuticals, Inc
305 Madison Avenue
Morristown, NJ 07960

Protocol Version and Date: Version 6.0, 18 Dec 2023

Analysis Plan Version: 3.0

Analysis Plan Date: 15 Oct 2025

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 09:37:22 EDT

13-Oct-2025
Date

[redacted] Biostatistics
Intercept Pharmaceuticals, Inc.

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 13, 2025 13:47:19 EDT

13-Oct-2025
Date

[redacted] Biostatistics
Intercept Pharmaceuticals, Inc.

[redacted]
Electronically signed by: [redacted]
Reason: I approve this document
Date: Oct 15, 2025 08:43:32 CDT

15-Oct-2025
Date

[redacted] Clinical Research
Intercept Pharmaceuticals, Inc.

TABLE OF CONTENTS

[redacted]

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or Specialist Term	Explanation
AE	adverse event
AESI	adverse event 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
CK	creatine kinase
CMH	Cochran-Mantel-Haenszel
CP	Child-Pugh
CS	clinically significant
CSR	clinical study report
DB	double-blind
DMC	Data Monitoring Committee
DOB	date of birth
eCRF	electronic case report form
ECG	electrocardiogram
[redacted]	
GGT	gamma-glutamyl transferase
ICH	International Conference on Harmonisation
INR	international normalized ratio

Abbreviation or Specialist Term	Explanation
IR	immediate-release
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]	
Na	sodium
NCS	not clinically significant
[redacted]	
OCA	obeticholic acid
OPTN	Organ Procurement and Transplantation Network
PBC	primary biliary cholangitis
PD	pharmacodynamic
[redacted]	
PK	pharmacokinetic
[redacted]	
PT	preferred term
QD	once daily
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SEM	standard error of the mean
SI	standard international unit

Abbreviation or Specialist Term	Explanation
SMQ	Standardized MedDRA Query
SOC	system organ class
SE	standard error
[redacted]	
TEAE	treatment-emergent adverse event
UDCA	ursodeoxycholic acid
ULN	upper limit of normal
US	United States
[redacted]	
WHO	World Health Organization

1. SCOPE

This statistical analysis plan (SAP) describes the statistical analyses and data presentations planned for protocol 747-214, which includes the Double-Blind (DB) Phase and Long-Term Safety Extension (LTSE) Phase. It provides a detailed description of the strategies, rationales, and statistical techniques to be used to meet the study objectives, as well as 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 before 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 plus BZF versus BZF alone in subjects with PBC on the following:

- Biochemical disease markers, including ALP, GGT, ALT, AST, total and conjugated bilirubin, and lipid panel
- Biomarkers of bile acid synthesis and homeostasis, including C4 and bile acids
- Safety and tolerability

2.1.3. Additional Objectives

[redacted]

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 100 mg vs BZF 100 mg and OCA + BZF 400 mg vs BZF 400 mg) with 2-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 group of the combination of OCA + BZF
- μ_b represents the mean change of ALP from baseline to Week 12 in the group of BZF

3. SUMMARY OF STUDY DESIGN

3.1. Overall Study Design

This is a Phase 2a, randomized, DB, active-controlled, parallel-group study in subjects with PBC as defined in Section 5 of the protocol. The study was designed to support the decision making for a future Phase 3 study.

This study will evaluate the efficacy, safety, and tolerability of OCA (5 mg/day) administered in combination with 2 different BZF doses (100 and 400 mg/day) compared to BZF alone (at 2 doses, 100 and 400 mg/day) in up to 60 subjects with PBC over at least 12 weeks. A rolling interim analysis was performed during the DB Phase to guide decision-making for the Phase 3 trial. No futility or superiority stopping rules will apply for the interim analysis. The Data Monitoring Committee (DMC) will perform an interim safety analysis when approximately 32 subjects complete the Week 12 visit in accordance with the DMC Charter. During 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 will be randomized (1:1:1:1) to the following 4 treatments: Treatment A: BZF 100 mg immediate-release (IR); Treatment B: BZF 400 mg IR; Treatment C: OCA 5 mg + BZF 100 mg IR; and Treatment D: OCA 5 mg + BZF 400 mg IR. This study will use 100-mg and 200-mg IR formulation tablets of BZF.

Subjects who are randomized to Treatment A or B will receive 100 or 400 mg, respectively, of BZF alone and those randomized to Treatment C or D (combination groups) will receive OCA 5 mg once daily and either 100 or 400 mg, respectively, of BZF. The DB Phase will be from Day 1 through Week 12; eligible subjects will then continue into the LTSE Phase through the end of the study.

After the Day 1 Visit, subsequent study visits during the DB Phase will occur at approximately Weeks 2 (±5 days), 4 (±5 days), 6 (±5 days), 8 (±5 days), 10 (±5 days), and 12 (±2 weeks) (end of DB Phase/Day 1 of LTSE) for the assessment of efficacy, safety, tolerability, and PK. These visits will also allow for the assessment of any adverse events (AEs), changes to concomitant medications and/or new medications that have been initiated, medical/surgical procedures, and to verify that the subject is dosing as directed. A rolling evaluation of available efficacy and safety data may occur periodically during the DB Phase.

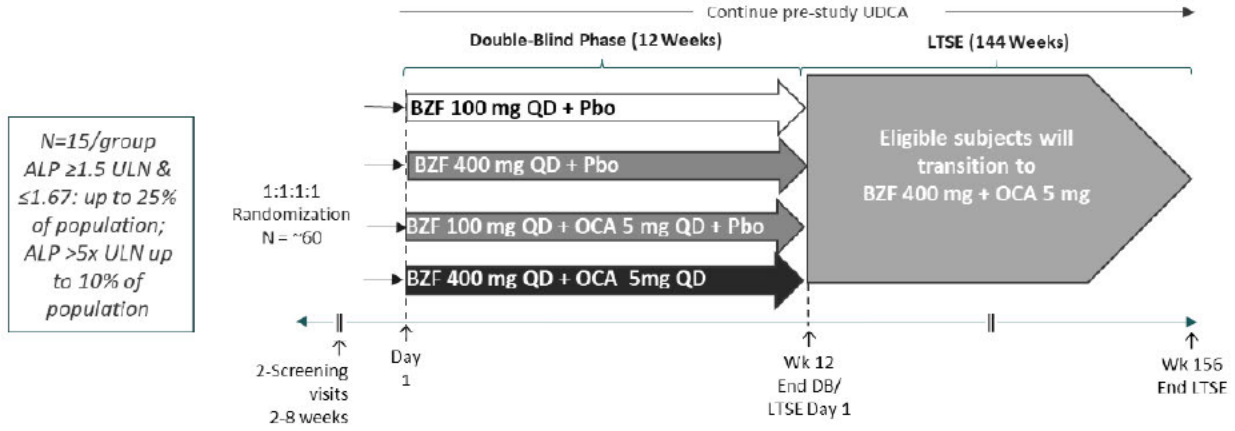
Subjects who continue to meet protocol requirements will transition directly to the LTSE Phase upon completion of the Week 12 of DB Phase and will receive treatment with OCA 5 mg once daily + BZF 400 mg once daily (Treatment D of the DB Phase) for the remainder of the study. All site staff will maintain study blind in DB Phase.

An independent DMC will review safety 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 charter.

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

3.2. Study Design Diagram

Figure 1: Study Design for the Double-Blind and LTSE Treatment Phases (Study 747-214)



ALP = alkaline phosphatase; BZF = bezafibrate; DB = double blind; LTSE = long-term safety extension; N = number of subjects; OCA = obeticholic acid; QD = once daily; UDCA = ursodeoxycholic acid; ULN = upper limit of normal; Wk=week
Note: First 3 subjects per treatment group will be assigned to PM dosing.
Note: Placebo = either OCA or BZF tablets
Note: Subjects taking UDCA at the time of enrollment will remain on their stable dose of UDCA during the study.
Note: LTSE Day 1 will be the last visit in the DB Phase (Week 12)
Note: Subjects who transition to the LTSE Phase will receive OCA 5 mg + BZF 400 mg IR QD.

3.3. Randomization and Blinding

This study will be conducted in a double-blind, randomized manner. Enrolled subjects will be randomized in a 1:1:1:1 ratio to one of the following treatment groups:

Treatment A: BZF 100 mg IR tablet once daily	OCA placebo tablet BZF 100 mg IR tablet BZF placebo tablet
Treatment B: BZF 400 mg IR tablet once daily	OCA placebo tablet BZF 200 mg IR tablet BZF 200 mg IR tablet
Treatment C: OCA 5 mg once daily + BZF 100 mg IR once daily	OCA 5 mg tablet BZF 100 mg IR tablet BZF placebo tablet
Treatment D: OCA 5 mg once daily + BZF 400 mg IR once daily	OCA 5 mg tablet BZF 200 mg IR tablet BZF 200 mg IR tablet

All subjects will be assigned IP 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) and

- ALP ≥1.5x ULN but ≤1.67x ULN (not more than 25% of subjects enrolled)
- ALP >1.67x ULN (remaining subjects enrolled)

Note: The number of subjects with baseline ALP ≥1.5x ULN but ≤1.67x ULN will not exceed 25% of subjects enrolled in the study and will include up to 10% of subjects with ALP >5x ULN.

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 of ALP in BZF is -60 U/L, 15 subjects per arm will provide 80% power to detect a treatment difference of 60 U/L between BZF arm and BZF+OCA arm at 2-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. 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 1 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 treatment actually received.

4.4. The Pharmacokinetic (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 Phase. The LTSE Population will be used for LTSE analyses.

5. GENERAL CONSIDERATION

Individual subject data obtained from 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. 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. CIs will be provided and will be rounded to 1 decimal place, unless otherwise specified, in the table and listing shell.

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

5.1. Baseline Definitions

The Baseline value for statistical analyses in the 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 the LTSE Population is defined as the last assessment on or before the date of transition to the LTSE Phase.

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.

Means, medians, SEMs, and SDs will be presented with an increased level of precision, where means and medians will be presented to 1 more decimal place than the raw data, and the SEMs and SDs will be presented to 2 more decimal places than the raw data. In general, the decimal places should not exceed 3 decimal places, unless appropriate.

For tables where rounding is required, rounding will be done to the nearest round-off unit. For example, when rounding to the nearest integer, values $\geq$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. 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%."

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 1 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 1 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 Phase:

- Randomized
- Treated (Safety)
- mITT 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 Phase:

- 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 the 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 Multiple Races)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported, Unknown, Multiply Ethnicities)
- 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 characteristic 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 (PTs) and system organ classes (SOCs) using the Medical Dictionary for Regulatory Activities (MedDRA).

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

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 DB Phase are expected to continue with their original scheduled visits until the end of the DB Phase 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 responder endpoints analyzed by Cochran-Mantel-Haenszel (CMH) test, 2 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's 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 Phase. 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 Double-blind Phase. For the primary efficacy analysis, the subjects with baseline and at least 1 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 computational algorithm fails to converge, the following structures will be tested in the order: Toeplitz, first-order autoregressive (AR [1]) and compound symmetry. If the model still cannot converge, the stratification factor of ALP (≥1.5x ULN but ≤1.67x ULN or >1.67x ULN) will be dropped from the model. If the model still fails to converge, the stratification factor of total bilirubin levels (≤0.7x or >0.7x ULN but <2x ULN) will be dropped from the model. 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 form 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 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]

Unless specified, the analysis methods for additional efficacy endpoints will be the same as those described for secondary efficacy endpoints.

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. CP 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

The ascites and hepatic encephalopathy variables are based on the investigator's clinical assessment.

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

CP 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 Phase will be summarized by visits using the baseline value for the LTSE Population. No statistical inference will be done on those endpoints.

9.5. Interim Analyses and Data Monitoring Committee

9.5.1. Interim Analyses

The Sponsor will perform an unblinded rolling interim analysis of efficacy, PK, and safety data throughout the DB Phase.

The efficacy measures for the rolling interim analysis include:

- Primary Endpoint: The change in ALP from baseline to Week 12 in the DB Phase in the mITT Population
- Secondary Endpoints:
  - The ALP response rates of ≥10%, ≥20%, ≥30%, and ≥40% reduction from baseline at Week 12 and the normalization rates of ALP at Week 12 in ITT Population

- Percentage of subjects with ALP <1.67x ULN, total bilirubin ≤ULN, and ALP decrease of ≥15% in the DB Phase in the ITT Population
- The changes from baseline to Week 12 in GGT, ALT, AST, total bilirubin, conjugated bilirubin and pruritus [redacted] in ITT Population

In addition, analyses of drug exposure and safety assessments will be performed.

A small group of Sponsor Staff will also be unblinded. The details will be specified in Data Access Plan.

In addition to the routine safety review, the DMC will also review the interim analysis.

[redacted]

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

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 their separate charter.

9.6. Multiple Comparisons/Multiplicity

No adjustment on type I error will be performed in this study.

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, unless otherwise specified. The analyses in DB Phase and LTSE Phase 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. Percent compliance will be summarized separately for subjects who completed the study versus those who withdrew early so as to distinguish between those subjects who were compliant throughout the entirety of the study versus those who were compliant until they withdrew from the study.

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 PTs and 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) in the DB Phase if it meets 1 or more of the following criteria:

- An AE starting on or after the first study drug dose of Double-Blind Phase.
- An AE occurring prior to the first study drug dose that worsens (increase in grade) after the first study drug dose during Double-Blind Phase.
- 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 the Double-Blind Phase or after the first dose of LTSE Phase is not counted as a TEAE in the Double-Blind Phase.

An AE is defined as a TEAE in the LTSE Phase if it meets 1 or more of the following criteria:

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

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. (SAE)

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 (AESI)

The 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 Phases separately. Summaries that are displayed by SOC and PT will be ordered by descending order of incidence of SOC and PT within each SOC. 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 AESI.

Treatment-emergent pruritus, defined as any PT 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 Double-Blind and LTSE Phases 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 AE, and onset of the first severe treatment-emergent pruritus AE 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 disorder TEAEs and SAEs.

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

- Time to first onset of treatment emergent hepatic disorders
  - 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 disorder 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 PT of myalgia. Summaries will include a summary of the crude incidence of myalgia TEAEs and SAEs.

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

- Time to first onset of treatment emergent myalgia
  - 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 myalgia 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.

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 both conventional and standard international system of units by treatment group using descriptive statistics at Baseline and at each on-study evaluation. The change from Baseline will also be summarized for DB and LTSE Phases 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 U/L
- GGT 0-65 IU/L (male) 0-60 IU/L (female)
- Total Bilirubin 1.71-12.3 µmol/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)
- Creatine kinase (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 or 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 from the central or local labs. The subjects meeting Hy's law criteria (defined as AST or ALT ≥3x ULN together with total bilirubin ≥2x ULN at any timepoint) will be summarized by treatment and visit for DB and LTSE Phases separately.

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

All clinical laboratory data (including central and local) 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 DB and LTSE Phases separately.

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

10.5. Electrocardiograms (ECGs)

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.

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

10.6. Physical Exam

The results of physical exams will be summarized by visit for DB and LTSE Phases 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 and 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 DB and LTSE Phases 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 1 medication per ATC class and preferred name. At each level of subject summarization, a subject is counted once if he/she reported 1 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]

[redacted]

13. DEVIATIONS FROM PROTOCOL

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

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.

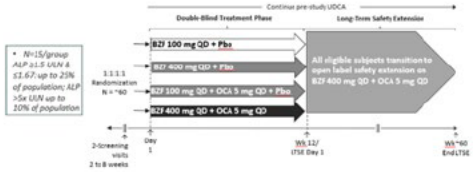
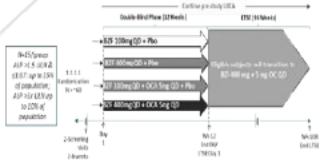
For categorical variables, the number and percentage of subjects within each category of the parameter, including a category for missing data, will be presented. Percentage calculations will be based on the analysis population, unless otherwise specified.

For continuous variables, summary statistics will include the number of subjects and the mean, standard deviation (SD), standard error of the mean (SEM), median, 25th and 75th 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:

The revisions from SAP Version 1.0 to SAP Version 2.0 summarized in this appendix include the following:

Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
[redacted]			
3.2	Study Design Diagram of the Double-Blind and LTSE Phases  …Note: LTSE Day 1 will be the last visit in the DB Phase (Week 12).	Study Design Diagram of the Double-Blind and LTSE Phases  …Note: LTSE Day 1 will be the last visit in the DB Phase (Week 12). **Note: Subjects who transition to the LTSE Phase will receive OCA 5 mg + BZF 400 mg QD. Once the optimized to-be marketed FDC dose is determined and agreed upon by the relevant regulatory authorities, subjects will receive the to-be-marketed FDC dose.**	Updated per protocol v5
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 **or 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 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
5.2	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.	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. **If imputed UDCA start date is earlier than the diagnosis date, the imputed UDCA start date will be set to the same as the diagnosis date.**	Updated imputation rule by including other possible scenarios
9	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 mITT population in Double-blind 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 Double-blind Period.	Updated to use ITT as the population for primary efficacy analysis to be consistent with the population used in other efficacy analyses.
[redacted]			

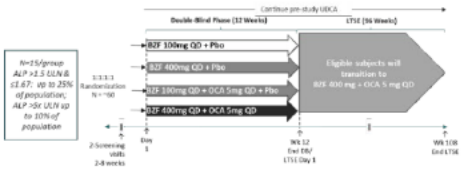
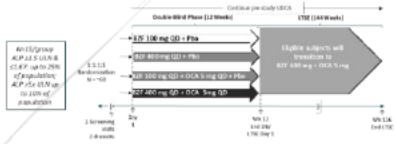
Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
[redacted]			
10.2.3	The AESIs include pruritus, hepatic injury and myalgia.	The AESIs include pruritus, hepatic injury, myalgia **and renal impairment**.	Updated per protocol v5
10.2.4.3	Myalgia are 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 disorder and investigational products, the following time-to event analyses will be performed:	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 **myalgia** and investigational products, the	Updated to fix the typos in version 1

Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
	• Time to first onset of treatment emergent hepatic disorder – 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 hepatic disorder will be censored at the date of last contact.	following time-to event analyses will be performed: • Time to first onset of treatment emergent **myalgia** – 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 **myalgia** will be censored at the date of last contact.	
10.2.4.4	NA	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. In order 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	Added analysis section per updated protocol v5

Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
		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 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.	
10.3	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	Laboratory parameters will be summarized in both conventional and standard international system of units by treatment group using descriptive statistics at Baseline and at each on-study evaluation	The analyses on conventional lab units are added per FDA's request.
10.3	Re-test or unscheduled results will not be summarized for DB and LTSE periods separately but will be included in the data listings.	Removed	The data from the local lab will also be included in the tables.
10.3	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 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)	The normal ranges of a number of lab parameters are updated. The data from the local lab will also be included in the tables.

Section	Original Text (Version 1.0)	Revised Text (Version 2.0)	Key Change Reasons/ Justification for Change
		• 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 or 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 from the central or local labs.	

The revisions from SAP Version 2.0 to SAP Version 3.0 summarized in this appendix include the following:

Section	Original Text (Version 2.0)	Revised Text (Version 3.0)	Key Change Reasons/ Justification for Change
[redacted]			
3.2	 Note: Subjects who transition to the LTSE Phase will receive OCA 5 mg + BZF 400 mg QD. ~~Once the optimized to-be-marketed FDC dose is determined and agreed upon by the relevant regulatory authorities, subjects will receive the to-be-marketed FDC dose.~~	 Note: Subjects who transition to the LTSE Phase will receive OCA 5 mg + BZF 400 mg **IR** QD.	Updated per protocol v6
5.1	The Baseline value for statistical analyses is defined as the last evaluations prior to the first administration of investigational product, unless otherwise specified.	The Baseline value for statistical analyses **in the 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 the LTSE Population is defined as the last assessment on or before the date of transition to the LTSE Phase.**	Updated the baseline definition

Section	Original Text (Version 2.0)	Revised Text (Version 3.0)	Key Change Reasons/ Justification for Change
7.1	• 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)	• Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White and **Multiple Races**) • Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported**, Unknown, Multiply Ethnicities**)	Updated per eCRF
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
[redacted]			

Section	Original Text (Version 2.0)	Revised Text (Version 3.0)	Key Change Reasons/ Justification for Change
[redacted]			
10.1	~~A summary of subjects who had an increase in dose and who had a decrease in dose at least once during the study will be provided in terms of frequency count (n) and percentages (%).~~	NA	Deleted the analysis
10.7	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 DB and LTSE Phases separately using the Safety and LTSE Populations.	Prior **and 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 DB and LTSE Phases separately using the Safety and LTSE Populations.	Updated the analysis