Statistical Analysis Plan for Protocol 747-207

A Phase 2, Randomized, Double-Blind, Placebo-Controlled, Dose-Finding, Clinical Trial Evaluating the Efficacy and Safety of Obeticholic Acid in Subjects with Primary Sclerosing Cholangitis

OBETICHOLIC ACID (OCA)

OBETICHOLIC ACID (OCA)

Protocol Version and Date:	Version 5: 18 March 2016
Phase:	Phase 2
Methodology:	Double-Blind, Randomized, Placebo-Controlled Study
Sponsor:	Intercept Pharmaceuticals, Inc. 4760 Eastgate Mall San Diego, CA 92121 Tel: 858-652-6800 Fax: 858-558-5961
Analysis Plan Date:	05 April 2017
Analysis Plan Version:	Final Version 2

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or Specialist Term	Explanation
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine transaminase
ANCOVA	analysis of covariance
Apo	apolipoprotein
ApoA1	apolipoprotein A-1
ApoB	apolipoprotein B
ApoE	apolipoprotein E
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC_{0-6}	area under the concentration-time curve from hour 0 to last sampling time (hour 6)
AUC_t	area under the concentration time curve
BLQ	below the limit of quantitation
BMI	body mass index
BP	blood pressure
BUN	blood urea nitrogen
C4	7α-hydroxy-4-cholesten-3-one
CA	cholic acid
CDAI	Crohn's disease activity index
CDCA	chenodeoxycholic acid
CI	confidence interval
CK-18	cytokeratin 18
cm	centimeter(s)
C_{max}	maximum plasma concentration
CSR	clinical study report
C_{ss}	steady state concentration
CV	coefficient of variation
DB	double-blind

Abbreviation or Specialist Term	Explanation
DCA	deoxycholic acid
DOB	date of birth
DOIC	date of informed consent
eCRFs	electronic case report forms
ECG	Electrocardiogram
eCRF	electronic case report form
ELF	enhanced liver fibrosis
eq	equivalents
ERCP	endoscopic retrograde cholangiopancreatography
FGF-19	fibroblast growth factor-19
FXR	farnesoid X receptor
g	gram(s)
GGT	gamma-glutamyl transferase
GI	Gastrointestinal
Glyco-CA	glycine 6α-ethyl cholic acid
Glyco-CDCA	glycine 6α-ethyl chenodeoxycholic acid
Glyco-DCA	glycine 6α-ethyl deoxycholic acid
Glyco-OCA	glycine 6α-ethyl chenodeoxycholic acid
Glyco-LCA	glycine 6α-ethyl lithocholic acid
Glyco-UDCA	glycine 6α-ethyl ursodeoxycholic acid
H	High
HA	hyaluronic acid
IBD	Inflammatory bowel disease
ICF	informed consent form
ICH	International Conference on Harmonization
IgA	Immunoglobulin A
IgG4	Immunoglobulin G4
IgM	Immunoglobulin M
IL	interleukin
INR	international normalized ratio
in	inch(es)
ITT	Intent-to-treat

Abbreviation or Specialist Term	Explanation
kg	kilogram(s)
KM	Kaplan-Meier
L	low
lb	pound(s)
LCA	lithocholate acid
LLN	lower limit(s) of normal
LLQ	lower limit of quantitation
ln	natural logarithm
LOCF	last observation carried forward
LS	least-square
LTSE	long-term safety extension
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram(s)
mL	milliliter
MRCP	magnetic resonance cholangiopancreatography
MW	molecular weight
ng	nanogram
NMR	nuclear magnetic resonance
OCA	obeticholic acid
ODS	output delivery system
P3NP	procollagen-3 N-terminal peptide
PD	pharmacodynamic
PK	pharmacokinetic(s)
PSC	primary sclerosing cholangitis
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SEM	standard error of mean
SI	international system of units
tauro-CA	taurine 6α-ethyl cholic acid
tauro-CDCA	taurine 6α-ethyl chenodeoxycholic acid

Abbreviation or Specialist Term	Explanation
tauro-DCA	taurine 6α-ethyl deoxycholic acid
tauro-LCA	taurine 6α-ethyl lithocholic acid
tauro-OCA	taurine 6α-ethyl chenodeoxycholic acid
tauro-UDCA	taurine 6α-ethyl ursodeoxycholic acid
TE	transient elastography
TEAE	treatment-emergent adverse event
TGF-β	transforming growth factor-beta
TIMP-1	tissue inhibitor of metalloproteinase 1
t_{max}	time to reach C_{max}
TNF-α	tumor necrosis factor alpha
UC	ulcerative colitis
UDCA	ursodeoxycholic acid
ULN	upper limit(s) of normal
VAS	visual analogue scale
W12C	Week 12 Completers
W24C	Week 24 Completers
WADD	Weighted average daily dose
WHODDE	World Health Organization Drug Dictionary Enhanced

1. INTRODUCTION

This document outlines the statistical methods to be implemented during the analyses of the double-blind (DB) phase data collected within the scope of Intercept Pharmaceuticals, Inc. (henceforth referred to as Intercept or the Sponsor) Protocol 747-207 (A Phase 2, Randomized, Double-Blind, Placebo-Controlled, Dose-Finding, Clinical Trial Evaluating the Efficacy and Safety of Obeticholic Acid in Subjects with Primary Sclerosing Cholangitis). The purpose of this plan is to provide specific guidelines from which the analyses will proceed. Any deviations from these guidelines will be documented in the clinical study report (CSR). The scope of this plan includes the detailed specifications of the statistical analyses for the DB phase only. A separate statistical analysis plan (SAP) will be written for the long-term safety extension (LTSE) phase of the study. The analyses described in this plan are considered a priori, in that they have been defined prior to database lock of the DB phase. Post hoc analyses will be labeled as such on the outputs and identified in the CSR. Further details about study design and procedures can be found in the protocol.

2. STUDY OBJECTIVES

The primary objective of the DB phase is to evaluate the effects of obeticholic acid (OCA) on the following in subjects with primary sclerosing cholangitis (PSC):

- Serum alkaline phosphatase (ALP)
- Safety

The secondary objectives are to evaluate the effects of OCA on the following in subjects with PSC:

- Hepatic biochemistry and indices of function
- Markers of:
  - Hepatic fibrosis and gastrointestinal (GI) inflammation and disease
  - Farnesoid X receptor (FXR) activity
  - Inflammatory bowel disease (IBD)
- Pharmacokinetics (PK) of OCA and Pharmacodynamic (PD) bile acids
- Exposure response of total OCA (OCA and its conjugates) to biomarkers (eg, ALP and bile acids)
- Long-term efficacy and safety of OCA
- Disease-specific symptoms

3. STUDY DESIGN AND PLAN

This is a Phase 2, randomized, double-blind, placebo-controlled, dose-finding evaluation of the efficacy and safety of OCA in subjects with PSC. The trial design is shown in Figure 1. Approximately 75 subjects who provide written informed consent and meet all of the inclusion

and none of the exclusion criteria will be randomized to 1 of 3 treatment groups as follows: 1.5 mg titrating to 3 mg OCA, 5 mg titrating to 10 mg OCA, or placebo in a 1:1:1 ratio. Subjects will administer investigational product orally, once daily for 2 consecutive 12-week periods. Randomization will be stratified by the presence or absence of concomitant UDCA use and total bilirubin level (≤1.5×ULN or >1.5×ULN but <2.5×ULN). No more than 50% of subjects randomized will be administering UDCA at the time of randomization.

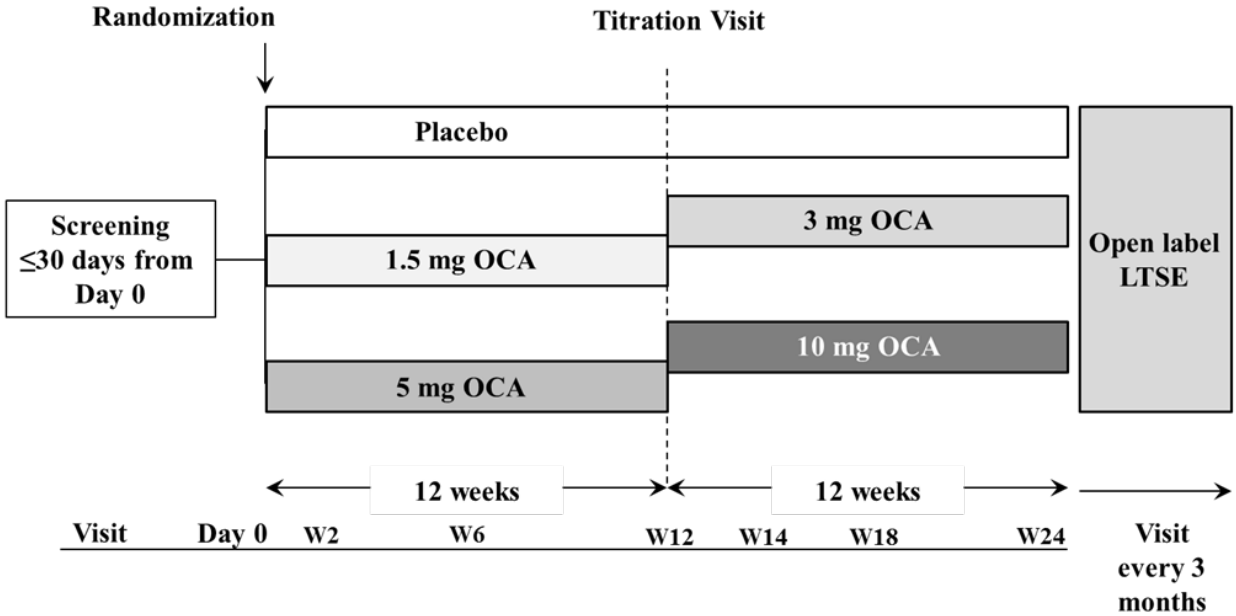
For the first 12 weeks after randomization, the subject's dose will be 1.5 mg OCA, 5 mg OCA, or placebo. After 12 weeks, the subject's dose will be titrated – 1.5 mg OCA titrating to 3 mg OCA and 5 mg OCA titrating to 10 mg OCA – and DB treatment will continue for a further 12 weeks at that dose.

Subjects will have a screening period of up to 30 days prior to Randomization/Day 0. Subjects will attend on site clinic visits at Weeks 2, 6, 12, 14, 18, and 24. The final visit during the DB phase will occur at Week 24, after which subjects will be asked to reconfirm their consent for participation in the LTSE phase (a further 24 months).

Upon a subject's completion of the DB phase, the trial blind will be broken in order to assign the starting OCA dose for the LTSE phase. It is intended that subjects will commence treatment at 5 mg OCA, except those subjects who completed treatment in the DB phase with 10 mg OCA who will continue at 10 mg OCA unless safety and tolerability warrant a dose reduction to 5 mg. The titration schedule and options for the LTSE phase are detailed in the protocol.

The overall study duration is up to 32 months, including up to a 30-day Screening period, 6 month DB period, followed by a 24-month open-label LTSE period, and a 1-month follow-up period. The LTSE may be extended as determined by the sponsor via protocol amendment and regulatory and Institutional Review Board/Independent Ethics Committee review and approval.

Figure 1: Study Design Schematic



4. DETERMINATION OF SAMPLE SIZE

A sample size of 25 subjects per treatment group, a total of 75 subjects, will provide at least 90% power to detect a treatment difference for change in ALP assuming 20% dropout and the mean absolute change in ALP for OCA and placebo treatment groups are approximately -20 and -5, respectively, with a pooled standard deviation of 16, based on a 2-sided independent 2-group t-test at an alpha level of 0.05.

In protocol Version 2, dated 24 Sep 2014, the sample size section was updated to reflect absolute change from percent change which was used in protocol Version 1, dated 06 June 2014. In the update, the percentage symbols were not removed. The percentage symbols have been removed in this section of the SAP, in order to accurately reflect what is used in the sample size calculation.

5. STUDY ENDPOINTS

5.1. Primary Endpoints

5.1.1. Primary Efficacy Endpoints

The primary efficacy endpoint is the Week 24 change from Baseline in ALP.

5.2. Secondary Endpoints

As defined in the protocol the secondary endpoints are ALP response rates, indices of hepatic function and GI inflammation, hepatic biochemistry, hepatic inflammation and fibrosis, PK, PD bile acids, and disease specific symptoms. The secondary efficacy endpoints of the DB phase are described below.

5.2.1. ALP Response Rates

The following ALP response rates will be summarized.

- ALP response, defined as ALP value <1.5 x upper limit of normal, will be evaluated at Baseline and all DB post-Baseline visits
- Percentage change from Baseline ≥10%, ≥15%, ≥20%, ≥25%, ≥30%, ≥35%, and ≥40% will be evaluated at all DB post-Baseline visits

5.2.2. Hepatic Biochemistry and Indices of Hepatic Function

The following laboratory parameters will be summarized for hepatic biochemistry and indices of hepatic function: ALP, albumin, alanine transaminase (ALT), aspartate aminotransferase (AST), total bilirubin, direct bilirubin, gamma-glutamyl transferase (GGT), international normalized ratio (INR).

- Observed values, change from Baseline, and percentage change from Baseline will be evaluated at each DB post-Baseline visit

5.2.3. FXR Activity

FXR activity will be assessed by measuring fibroblast growth factor-19 (FGF-19).

- Observed values, change from Baseline, and percentage change from Baseline will be evaluated at each DB post-Baseline visit

5.2.4. Disease Specific Symptoms

The disease specific endpoints include the following: Pruritus Visual Analogue Scale (VAS), 5-D Itch questionnaire, partial Mayo score assessment for subjects with ulcerative colitis (UC), and Crohn's disease activity index (CDAI).
Observed values and change from Baseline will be evaluated at each DB post-Baseline visit

-
-

5.3. Exploratory Endpoints

5.3.1. Exploratory Marker of Pruritus

Pruritus will be assessed by measuring Autotaxin activity.

- Observed values, change from Baseline, at Week 12 and Week 24, will be summarized
- Correlation of Autotaxin with 5D and VAS scores at Week 12 and Week 24 will be assessed

6. GENERAL ANALYSIS CONSIDERATIONS

6.1. Data Reporting

The statistical analyses will be reported using summary tables, figures, and data listings.

Laboratory units will be summarized and presented both in conventional units (CV) and in the international system of units (SI).

Individual subject data obtained from the electronic case report forms (eCRFs), external laboratory data, and any derived data (such as change from Baseline and percent change from Baseline) will be presented in data listings by subject. Data from all assessments, whether scheduled or unscheduled, will be listed by subject and visit. Unscheduled visits and visits occurring more than one day outside protocol defined window will not be included in the table summaries, excluding transient elastography measurements.

The analyses described in this plan are considered a priori, in that they have been defined prior to database lock of the DB phase. Post hoc analyses will be labeled as such on the outputs and identified in the CSR.

All analyses and tabulations will be performed using SAS® Version 9.2 or higher. PK parameters will be estimated using Phoenix® WinNonlin® Version 6.3. The following processes will be employed to validate statistical outputs: derived datasets (both tabulation datasets and analysis datasets), summary tables, and data listings will be verified through independent programming; graphical displays will be compared against supporting summary tables; and all outputs will undergo a senior-level statistical review. The process includes confirmation that statistically valid methods have been implemented and that all data manipulations and calculations are appropriate and accurate. Checks will be made to ensure accuracy, adherence to this SAP, consistency within tables, and agreement between tables and their corresponding data listings. Upon completion of validation/verification and quality review procedures, all documentation will be collected and filed in the study master file by the project statistician or designee.

6.2. Data Analysis and Summaries

Data distribution characteristics will determine which analysis methods are most appropriate. If methods do not allow for parametric modeling assumptions to be met, then non-parametric methods will be implemented.

6.2.1. Arithmetic Summaries

Continuous variables will be summarized by means, standard deviations (SDs), standard errors of the mean (SEMs), medians, interquartile range (IQR), minimums, and maximums.

6.2.2. Categorical Methods

Categorical variables will be summarized by counts and percentage of subjects in corresponding categories. Percentages will be based on the number of non-missing assessments unless otherwise specified.

6.2.3. Analysis of Covariance

Analysis of covariance (ANCOVA) will be performed to provide least square (LS) mean estimates and 95% two-sided confidence intervals (CIs) of the change and percentage change from Baseline at each DB post-Baseline visit. Model covariates include then randomization stratification values (ie, UDCA use [yes, no] and screening total bilirubin [≤1.5 x ULN, >1.5 x ULN – <2.5 x ULN]) and the Baseline of the parameter being estimated. The primary analysis will exclude the treatment by visit interaction term.

Normality/Equal Variance Testing

The assumption of a random error component being normally distributed will be tested by examining skewness and kurtosis following D'Agostino et al. (1990) with the residuals pooled over treatments. Additionally, homogeneity of variances across treatment groups will be examined using Levene's Test (Glaser 1982) with significance level of 0.05, based on absolute values of the residuals from the residual medians for each treatment. The data will be determined non-normal if the p-value for either the D'Agostino Skewness Test or D'Agostino Kurtosis Test is less than 0.05. If the data do not meet the normality or homogeneity assumptions of the parametric analyses, then a RT-2 rank transformation will be applied to the data (CFB endpoint and baseline value). To minimize the effect of differing study site sizes, the ranks will be standardized to lie between 0 and 1 using the NPLUS1 option of SAS® PROC RANK. Tied values will receive the mean value (midranks) of the corresponding ranks.

Example SAS® code for the primary analysis is as follows:

```
ODS graphics on;
proc mixed data=ADLB noclprint plots=residualpanel;
        class udcafl tbilfl trt;
        model change=Baseline udcabl tbilfl trt / ddfm=kr;
        lsmeans trt / cl diff alpha=0.05;
        title1 "Linear Mixed Effects Model Based Estimates of XXXX Means";
        title2 " Fixed Effects for Visit, Randomization Strata, and Treatment and Baseline as
                Covariates";
run;
ODS graphics off;
quit;
```

6.2.4. Analysis of Covariance (Repeated Measures)

Analyses of the clinical laboratory values will also be carried out using a restricted maximum likelihood (REML) based repeated measures linear mixed model (MMRM) to evaluate the effect over time by providing LS mean estimates, 95% CIs of change and percentage change from Baseline at each DB post-Baseline visit. Model covariates include then randomization stratification values (ie, UDCA use [yes, no] and pre-randomization total bilirubin [≤1.5 x ULN, >1.5 x ULN – <2.5 x ULN]), Baseline of the parameter being estimated, and a treatment by visit interaction term. An unstructured (5×5) covariance model will be used. If the computational algorithm fails to converge, the following structures will be executed: heterogeneous Toeplitz, Toeplitz, heterogeneous First-Order Autoregressive [AR (1)], heterogeneous compound symmetry (HCS), and compound symmetry (CS). The covariance structure converging to the best fit, as determined by Akaike's information criterion (AIC), will be used. The Kenward and Roger method will be used to calculate the denominator degrees of freedom for the test of fixed effects.

Example SAS® code for the ANCOVA with repeated measures is as follows:

```
ODS graphics on;
proc mixed data=ADLB noclprint plots=residualpanel;
        class subjid week ucdafl tbilfl trt;
        model change=week Baseline ucdafl tbilfl trt trt*week/ ddfm=kr;
        repeated week / subject=subjid type=uns;
        lsmeans trt trt*week / cl diff alpha=0.05;
        title1 "Linear Mixed Effects Model Based Estimates of XXXX Means by Visit";
        title2 " Fixed Effects for Visit, Randomization Strata and Baseline as Covariates, and
Treatment by Visit Interaction";
run;
ODS graphics off;
quit;
```

6.2.5. Median and Confidence Interval Estimation Methods

Hodges Lehmann estimation will be used to provide estimators of the median difference and 95% CIs for the median difference.

6.2.6. Geometric (Natural Log Transformed) Summaries

In order to calculate a geometric mean and corresponding 95% CI, the following steps are used:

- Transform the data by taking natural logarithms ($\ln = \log_e$)
- Calculate the mean and 95% CI of the $\log_e$-transformed data
- Exponentiate the mean and the lower and upper CIs back to the original scale in order to obtain the geometric mean and corresponding 95% CI.

The geometric coefficient of variation (geometric CV%) is calculated as $100*\text{sqrt}(e^{SD**2}-1)$ where SD is the SD of the $\log_e$-transformed data.

6.2.7. Multiple Comparisons/Multiplicity

A hierarchical approach will be used for multiplicity adjustments. If the primary efficacy analysis is statistically significant ($p < 0.05$), the following order will be used in the testing procedure to compare the change from Baseline in ALP between OCA and placebo:

- Week 12: OCA 5 mg treatment group (randomized to 5 mg for the initial 12 weeks followed by 10 mg for the latter 12 weeks) vs. placebo
- Week 24: OCA 3 mg treatment group (randomized to 1.5 mg for the initial 12 weeks followed by 3 mg for the latter 12 weeks) vs. placebo
- Week 12: OCA 1.5 mg treatment group (randomized to 1.5 mg for the initial 12 weeks followed by 3 mg for the latter 12 weeks) vs. placebo

If at any step the comparison is not statistically significant, then all subsequent comparisons will be exploratory rather than confirmatory.

Multiplicity adjustments will not be applied for any other statistical testing.

6.2.8. Subgroup Analyses

Subgroup analyses will be evaluated as deemed appropriate. Subgroups are defined as follows:

- Antibiotic use after first dose (Yes/No) – World Health Organization Anatomical Therapeutic Chemical (ATC) codes starting with ‘J01’and the medication start date is on or after the first dose of investigational product; for analysis of the primary efficacy endpoint.
- Age (≤65 years, >65 years),
- BMI (<30 kg/m^2 vs ≥30 kg/m^2)
- Baseline UDCA use (Yes vs No)
- Baseline Total Bilirubin (≤1.5xULN vs >1.5xULN)
- ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown

- race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White)
- gender (Male, Female)
- IBD (Yes vs No)
- Crohn's disease (Yes vs No)
- Ulcerative colitis (Yes vs No)
- Median TE Score (< vs ≥ Median score for all ITT subjects)

6.3. Data Handling

6.3.1. Baseline Values

Baseline values are defined as follows:

- **Efficacy, Clinical laboratory, and PD endpoints**: the mean of all available evaluations prior to first dose of investigational product.
- **Lipoprotein assessments**: the last non-missing fasted assessment prior to first dose of investigational product.
- **ECG assessment**: the last available non-missing assessment prior to first dose of investigational product.

For any other quantitative parameters, Baseline is defined as the mean of all available evaluations prior to administration of investigational product on Day 1.

6.3.2. Missing Data

For secondary efficacy analyses of ALP response rates, missing values will be considered as non-responders.

Additional sensitivity analysis will be performed using LOCF, and it will also be performed modifying the handling missing data for ALP response criteria. For the modified response criteria, all OCA treated subjects, missing data will be considered non-responders. For Placebo patients, missing data will be considered responders.

Otherwise, analysis will be done using observed data only; missing values will not be imputed.

6.3.3. Partial Dates

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

- Diagnosis date (eg, PSC diagnosis date)
  - For missing day only: Day will be imputed as the first day of the month (ie, 01).
  - For missing day and month: Day and month will be imputed as the first day of the year (ie, 01 January).
- Start dates (eg, AE onset date or start date of medication)

  - For missing start day only: Day will be imputed as the first day of the month (ie, 01) with the following exception: if the partial date falls in the same month and year as the date being used in the calculation (eg, first dose date, informed consent date), then the partial date will be imputed to equal the date being used for the calculation.
  - For missing start day and month: Day and month will be imputed as the first day of the year (ie, 01 January) with the following exception: if the partial date falls in the same year as the date being used in the calculation (eg, first dose date, informed consent date), then the partial date will be imputed to equal the date being used for the calculation.
  - Imputed start dates must be prior to the stop date.
- Stop dates (eg, AE resolution date or stop date of medication)
  - For missing stop day only: Day will be imputed as the last day of the month (ie, 28, 29, 30, or 31).
  - For missing stop day and month: Day and month will be imputed as the last day of the year (ie, 31 December).
  - Imputed dates should not extend beyond the DB phase (ie, date of completion or discontinuation of the DB phase).
  - Imputed stop dates must be on or after the start date.

6.3.4. Data Conventions

Unless otherwise specified, in summary tables of continuous variables, the minimum and maximum values will be displayed to the same number of decimal places as the raw data, the mean and median will be presented to one extra decimal place compared to the raw data, and the SD and SE of the mean will be displayed to two extra decimal places compared to the raw data. Rounding will only occur after all calculations have been incorporated. For tables where rounding is required, rounding will be done to the nearest round-off unit; for example, when rounding to the nearest integer, values ≥XX.5 will be rounded up to XX + 1 (eg, 97.5 will round up to 98), whereas values <XX.5 will be rounded down to XX (eg, 97.4 will round down to 97).

Percentages based on frequency counts will be based on available data, and denominators will generally exclude missing values unless some form of imputation is defined or a 'missing' category is presented. For frequency counts of categorical variables, categories with zero counts 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%."

6.3.5. Standard Calculations

Variables requiring calculation will be derived using the following formulas:

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

  duration (days) = date2 – date1 + 1

- **Months**: A duration expressed in months is calculated as the number of days divided by 365.25 / 12.

- **Years**—A duration expressed in years between one date (*date1*) and another date (*date2*) is calculated using the following formulas:

  duration (years) = (date2 – date1+ 1) / 365.25

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

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

  Otherwise, the following formula is used:

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

- **Height**: Height measured in inches (in) are converted to centimeters (cm) using the following formula:

  height (cm) = height (in) × 2.54

- **Weight**: Weight measured in pounds (lb) are converted to kilograms (kg) using the following formula:

  weight (kg) = weight (lb) / 2.2046

- **Temperature**: Temperature measured in degrees Fahrenheit are converted to degrees centigrade using the following formula:

  temp (degrees centigrade) = 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 is calculated as:

  change from Baseline = DB post-Baseline value – Baseline value

- **Percentage change from Baseline**: Percentage change from Baseline is calculated as:

  percentage change from Baseline = 100 × ([DB post-Baseline value – Baseline value] / Baseline value)

- **Geometric mean -** In order to calculate a geometric mean and corresponding 95% CI, the following steps are used:

  – Transform the data by taking natural logarithms ($\log_e$)

- Calculate the mean and 95% CI of the $\log_e$-transformed data
- Exponentiate the mean and the lower and upper CIs back to the original scale in order to obtain the geometric mean and corresponding 95% CI.

The geometric coefficient of variation (geometric CV%) is calculated as 100*sqrt(e^{SD**2}-1) where SD is the SD of the $\log_e$-transformed data.

- **Coefficient of variation** - The coefficient of variation (%) will be calculated as the ratio of the standard deviation to the arithmetic mean using the following formula:

$$\text{Coefficient of variation} = \frac{\text{standard deviation}}{\text{arithmetic mean}}$$

6.3.6. Pruritus 5-D Itch Score Calculations

The 5-D questionnaire (Elman 2010) is a specific tool used to quantify the magnitude of pruritus in PBC (and other diseases). It consists of 5 domains: duration, degree, direction, disability, distribution, as well as the total score. The following rules will be used to calculate the domain scores:

- Single item domain scores (duration, degree, and direction) are equal to the value indicated below the response choice (range from 1-5).
- The disability domain includes 4 items that assess the impact of itching on daily activities: sleep, leisure/social activities, housework/errands, and work/school. The score for the disability domain is achieved by taking the highest score on any of the 4 items. The disability domain will only be calculated if at least 3 daily activities are documented, otherwise this domain score will be set to missing.
- For the distribution domain, only the section "Mark whether itching has been present in the following parts of your body over the last 2 weeks" will be used. The distribution domain includes 16 potential locations of itch, including 15 body part items and one point of contact with clothing. Points of contact with clothing will only count as one point. Missing ticks are interpreted to be absent. The number of affected body parts ('present') is tallied (potential sum 0 to 16) and the sum is sorted into 5 scoring bins: sum of 0 to 2 = score of 1, sum of 3 to 5 = score of 2, sum of 6 to 10 = score of 3, sum of 11 to 13 = score of 4, sum of 14 to 16 = score of 5.
- The total 5D score is obtained by summing up the domain scores and ranges between 5 (no pruritus) and 25 (most severe pruritus). The total score will not be calculated if any of the domain scores is missing.
- If multiple answers are given for any item then the most severe one will be used.

7. ANALYSIS POPULATIONS

The following analysis populations will be used:

7.1. Intent-to-Treat (ITT) Population

All randomized subjects who receive any amount of investigational product will be included in the ITT population. Treatment assignment will be based on the randomized treatment. The ITT population will be used for the analysis of all efficacy data.

7.2. Week 12 Completer Population

The Week 12 Completer (W12C) population will include all ITT subjects who complete the DB Phase Week 12 ALP assessment and do not have a major protocol deviations that potentially affect the efficacy of the study drug. Treatment assignment will be based on the randomized treatment.

7.3. Week 24 Completer Population

The Week 24 Completer (W24C) population will include all ITT subjects who complete the DB Phase Week 24 ALP assessment and do not have a major protocol deviations that potentially affect the efficacy of the study drug. Treatment assignment will be based on the randomized treatment.

7.4. Per Protocol Population

All randomized subjects without major protocol deviations that potentially affect the efficacy of the study drug will be included in the per protocol population. Subjects who are in the study through Week 12 who do not up titrate their dose of IP will be excluded from the per protocol population. Treatment assignment will be based on the randomized treatment.

7.5. Safety Population

The Safety population will include all subjects who receive any amount of investigational product. Treatment assignment will be based on the treatment actually received. The Safety population will be used for the analysis of all safety data.

8. STUDY POPULATION

8.1. Subject Disposition

Subject disposition information will be summarized and listed for all subjects. Summaries will include the following: the number of subjects randomized, the number of subjects in each analysis population, the number of subjects at each center, the number of subjects completing the DB phase (Week 24) and each scheduled visit (Week 2, 6, 12, 14, and 18), the primary reason for discontinuation from the DB phase, and the number of subjects enrolling in the LTSE. A subject will be considered as completing the DB phase (Week 24) and each scheduled visit (Weeks 2, 6, 12, 14, and 18) if the subject has a non-missing assessment of ALP at the corresponding visit. Percentages will be based on the ITT Population.

8.2. Protocol Deviations

Protocol deviations for missed visits, missed assessments, out of window visits or assessments, and violations of inclusion/exclusion criteria will be determined based on available data. All other protocol deviations will be collected by the clinical research associates. Major protocol deviations identified during the DB phase that could potentially affect the conclusions of the study or result in a subject's removal from an analysis population will be classified as such prior to database lock of the DB phase. Major protocol deviations will be summarized by deviation category and treatment group. All protocol deviations and their classifications will be presented in a listing. Subjects with protocol deviations that result in a subject's removal from the ITT, Safety, W12C, W24C, Per Protocol, or PK Populations will be flagged in a listing.

8.3. Demographic and Baseline Characteristics

Demographic variables will include the following:

- Age at informed consent
- Age at informed consent categorized as <65 years and ≥65 years
- Sex
- Race/Ethnicity

Other Baseline characteristics will include the following:

- Weight (kg)
- Height (cm)
- BMI (kg/m^2)
- BMI categorized as <30 kg/m^2 and ≥30 kg/m^2
- Hepatic ultrasound to assess bile duct patency performed (Yes/No)
- Key liver function test results as a continuous variable and categorized as follows by the ULN and the LLN unless otherwise specified:
  - ALP: ≤ 3xULN, >3xULN; and, ≤ median baseline ALP, > median baseline ALP

  - Total Bilirubin: ≤ ULN, > ULN, > ULN - ≤1.5xULN, >1.5xULN - ≤2xULN, >2xULN - ≤2.5xULN
  - International normalized ratio (INR): ≤1.3, >1.3 stratified by subjects on/not on anticoagulants at the time of randomization

Arithmetic summary statistics as described in Section 6.2.1 will be presented for age, weight, height, BMI, and key liver function test results. Frequency counts and percentages will be presented for age groups, sex, race/ethnicity, BMI category, key liver function categories, and hepatic ultrasound.

Race and ethnicity will be summarized as follows:

- Race
  - American Indian or Alaska Native
  - Asian
  - Black or African American
  - Native Hawaiian or Other Pacific Islander
  - White
- Ethnicity
  - Hispanic or Latino
  - Not Hispanic or Latino
  - Not reported
  - Unknown

Demographic and Baseline characteristics will be summarized for the ITT, Safety, PK, W12C, and W24C populations. All demographic and Baseline characteristics will be presented in data listings. The subgroup tables will be generated for IBD, UC, and CD regardless of the total number of patients in each group. For the remaining subgroups, the tables will be generated as long as each subgroup contains at least 10% of the subjects in the overall population.

8.4. PSC Disease History

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

- Age at first PSC diagnosis
- Age at first PSC diagnosis categorized as <40 years and ≥40 years
- Symptomatic at diagnosis (Yes/No)
- Age at first occurrence of PSC symptoms
- Age at first occurrence of PSC symptoms categorized as <40 years and ≥40 years
- Duration of PSC in years at time of informed consent

- Duration of PSC categorized as ≤ median years and > median years
- PSC diagnosis confirmed via cholangiagraphy?
- Modality of PSC diagnosis (endoscopic retrograde cholangiopancreatography [ERCP], magnetic resonance cholangiopancreatography [MRCP], Other)
- Affected bile ducts (intra-hepatic, extra-hepatic, both, unknown)
- Symptoms of PSC (Never, Previous, Current)
  - Right upper quadrant abdominal pain
  - Fatigue
  - Weight loss
  - Jaundice
  - Pruritus
  - Other
- Severity of most recent pruritus event (Mild, Moderate, Severe)
- UDCA Use (Never/Previous/Current)
- Total daily dose of UDCA (mg/kg)
- Inflammatory Bowel Disease (IBD) (No IBD, Ulcerative Colitis, Crohn's Disease, Other)

Arithmetic summary statistics as described in Section 6.2.1 will be presented for the age at first PSC diagnosis, age at first occurrence of PSC symptoms, duration of PSC, and total daily dose of UDCA (mg/kg). All other categorical PSC disease characteristics will be summarized using frequency counts and percentages. PSC disease history will be summarized for the ITT, W12C, and W24C Populations. The subgroup tables will be generated for IBD, UC, and CD regardless of the total number of patients in each group. For the remaining subgroups, the tables will be generated as long as each subgroup contains at least 10% of the subjects in the overall population. PSC disease history will be presented in a data listing.

8.5. General Medical History

Verbatim medical history terms on eCRFs will be mapped to preferred terms and system organ classes using the Medical Dictionary for Regulatory Activities (MedDRA) version 17.1. Frequency counts and percentages of the number of subjects reporting an abnormal Baseline medical history will be summarized by MedDRA system organ class and preferred term using the ITT Population.

Summaries that are displayed by system organ class and preferred term will be ordered by descending order of incidence of system organ class and preferred terms within each system organ class.

Medical history data will be presented in a data listing.

8.6. Prior and New Concomitant Medications

Verbatim terms on eCRFs will be mapped to ATC class and preferred term using the World Health Organization Drug Dictionary Enhanced (WHO-DDE June 2014).

Pretreatment medications are those medications with start and stop dates prior to the first dose of investigational product in the DB phase. Prior concomitant medications are those medications that started prior to, and continued after, the first dose of investigational product in the DB phase. New concomitant medications are those medications that were started after the first dose of investigational product in the DB phase. If it cannot be determined whether the medication was a new concomitant medication due to a partial start or stop date or if the medication is taken on the same date as the first OCA dose in the DB phase, then it will be counted as a new concomitant medication.

Pretreatment medications will be presented in listings only. Prior and new concomitant medications will be summarized by World Health Organization ATC class and preferred term using the ITT Population. New concomitant medications will be summarized separately. 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. A subject is counted once if one or more medication is reported at the ATC class or preferred term level. Each summary will be ordered by descending order of incidence, for all OCA treated patients, of ATC class and preferred term within each ATC class.

Prior and new concomitant medications will be presented a data listing.

9. EFFICACY ANALYSES

The primary and secondary efficacy analyses will be based on the ITT Population, W12C and W24C populations.

9.1. Primary Efficacy Analyses

The primary efficacy endpoint is the Week 24 change from Baseline in ALP. Using the ITT population, the primary efficacy analysis will compare the Week 24 change from Baseline in ALP between OCA 10 mg treatment group and placebo and OCA 3 mg treatment group and placebo using an ANCOVA model with treatment group, visit, and randomization strata as fixed effects, and Baseline ALP as a covariate.

In order, to provide estimates of the change from Baseline and corresponding 95% CI, the repeated measures ANCOVA methods described in Section 6.2.3 will be applied. Model fit will be assessed in order to determine the validity of the estimates. The analysis will be repeated with percentage change from Baseline as the dependent variable. Estimates of least-square (LS) means, SEs, and 95% CIs will be presented by treatment group.

The values, change from Baseline, and percentage change from Baseline will be summarized using observed values by treatment group and visit using arithmetic summary statistics as described in Section 6.2.1.

A hierarchical approach will be used for multiplicity adjustment across the primary efficacy endpoint for other visits and doses as described below. If the primary efficacy analysis is statistically significant ($p < 0.05$), the following order will be used in the testing procedure to compare the change from Baseline in ALP between OCA and placebo:

- Week 12: OCA 5 mg treatment group (randomized to 5 mg for the initial 12 weeks followed by 10 mg for the latter 12 weeks) vs. placebo
- Week 24: OCA 3 mg treatment group (randomized to 1.5 mg for the initial 12 weeks followed by 3 mg for the latter 12 weeks) vs. placebo
- Week 12: OCA 1.5 mg treatment group (randomized to 1.5 mg for the initial 12 weeks followed by 3 mg for the latter 12 weeks) vs. placebo

If at any step a comparison above is not statistically significant, then all subsequent comparisons will be exploratory rather than confirmatory.

Sensitivity analysis for the primary efficacy analysis will be performed using the W12C and W24C Populations. In addition, the arithmetic summary statistics will be presented by antibiotic use, IBD, UC, and CD subgroups as described in Section 6.2.8.

An unblinded interim analysis of the primary endpoint will be conducted after approximately 50% of the subjects have completed the initial 12-weeks of blinded treatment. The interim analysis will use the same model described in Section 6.2.3 on the change from Baseline ALP to Week 12 in addition to Week 24.

9.2. Secondary Efficacy Analyses

Secondary efficacy analyses of hepatic biochemistry and function parameters as described in Section 5.2.2 will be summarized by treatment group using descriptive statistics at Baseline and at each scheduled DB post-Baseline visit. The change from Baseline will also be summarized. Hepatic biochemistry and function parameters will be analyzed using the same ANCOVA model as specified for the primary efficacy analysis.

9.2.1. ALP Responders

ALP response rates, defined as ALP to <1.5×ULN, will compare OCA treatment groups vs. placebo at Baseline and all DB post-baseline visits using a Cochran-Mantel-Haenszel test stratified by the randomization stratification factor. For this analysis, missing values will be imputed as described in Section 3.2.1.

The proportion of responders will be summarized by treatment group and visit as described in Section 6.2.2 for W24C, W12C, and ITT populations.

9.2.2. Hepatic Biochemistry and Indices of Function

The secondary efficacy parameters related to hepatic biochemistry and are described in Section 5.2.2.

These parameters will be analyzed using arithmetic summary statistics, ANCOVA, and model estimates with independent variables as described Section 6.2.3 for the primary efficacy endpoint.

Calprotectin will potentially be analyzed. The details of the analysis will be specified in a separate SAP.

9.2.4. Disease-specific symptoms

The Pruritus VAS, Pruritus 5-D Itch questionnaire, Partial Mayo scoring system for assessment of UC activity (partial Mayo score), and Crohn's disease activity index (CDAI) will be compared between OCA treatment groups and placebo at Week 12 and Week 24 using a Wilcoxon rank-sum test. This analysis will be repeated for change from Baseline and percent change from Baseline in these scores.

The proportion of subjects in remission from Crohn's disease (CDAI < 150) will be summarized by treatment group and visit using statistics as described in Section 6.2.2, and compared between treatment groups using Fisher's exact test. The analysis will be repeated for the proportions of subjects in remission from UC (partial Mayo score of ≤2 with no individual sub-score exceeding 1), and the proportion with mild disease (a partial Mayo score ≤3 with no individual sub-score exceeding 1 point).

9.3. Exploratory Analyses

9.3.1. FXR Activity

FXR activity will be assessed by measuring FGF-19. This parameter will be analyzed using arithmetic summary statistics, ANCOVA, and model estimates with independent variables as described in Section 6.2.3 for the primary efficacy endpoint.

10. SAFETY ANALYSES

All safety analyses will be based on the Safety Population. The evaluation of the effects of OCA on safety parameters in the PSC population is a primary objective of this study.

10.1. Extent of Exposure

10.1.1. Investigational Product (OCA or Placebo)

Subjects will titrate their dose of OCA or placebo at the Week 12 visit according to the criteria specified in the protocol. Frequency counts and percentages of the number of subjects who do not up-titrate will be summarized. The duration of investigational product exposure 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 investigational product (mg) exposed to subject will be calculated by adding the doses taken by a subject during the study and will be summarized using descriptive statistics.

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

- (# of days consumed during study)/ (# of days on study drug, excluding drug holidays) * 100

Investigational product compliance will be summarized by treatment group using descriptive statistics.

Additional summaries will present the number and percentage of subjects with any drug interruption, subjects with alternate dosing, subjects with reduced dose, subjects who did not titrate dose at Week 12, and number of subjects with permanent discontinuation. Denominators for calculating percentages will be based on the number of subjects who received at least one

dose in the treatment group summarized. All exposure data will be presented in a subject data listing.

10.2. Adverse Events (AEs)

All AE summaries will be restricted to treatment-emergent adverse events (TEAEs), which are defined as any AEs that newly appear, increase in frequency, or worsen in severity following initiation of study medication. If it cannot be determined whether the AE is treatment emergent due to a partial onset date, then it will be counted as treatment emergent. Verbatim terms on eCRFs will be mapped to preferred terms and system organ classes using MedDRA (version 17.1).

Summaries that are displayed by system organ class and preferred terms will be ordered by descending order of incidence for all OCA treated patients of system organ class and by preferred terms within each system organ class. Summaries of the following types will be presented by treatment group:

- Overall summary of TEAEs
- Subject incidence of TEAEs and the total number of entries by MedDRA system organ class and preferred term.
- Subject incidence of drug-related TEAEs by MedDRA system organ class and preferred term. Related AEs are those with relationships reported as “Definite,” “Probable,” “Possible,” or with a missing relationship.
- Subject incidence of serious adverse events (SAEs) by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs leading to investigational product withdrawal or study discontinuation by MedDRA system organ class and preferred term. This is a subset of the AEs where Action Taken with Study Medication is checked as “Drug Withdrawn” or Subject Discontinued from Study is checked.
- Subject incidence of TEAEs leading to investigational product withdrawal by MedDRA system organ class and preferred term. This is a subset of the AEs where Action Taken with Study Medication is checked as “Drug Withdrawn”.
- Subject incidence of TEAEs leading to study discontinuation by MedDRA system organ class and preferred term. This is a subset of the AEs where Subject Discontinued from Study is checked.
- Subject incidence of TEAEs by MedDRA system organ class, preferred term, and maximum severity will be presented. At each level of subject summarization, a subject is classified according to the maximum severity if the subject reported one or more events. AEs with missing severity will be considered severe for this summary.
- Subject incidence of follow-up TEAEs and the total number of entries by MedDRA system organ class and preferred term. Follow-up TEAEs are defined as those TEAEs with onset after the Week 24 visit, for those subjects not continuing to the LTSE phase.

The summaries above will be repeated presenting only preferred term ordered by descending order of incidence for all OCA treated patients, excluding system organ class.

The incidence of pre-treatment AEs and pre-treatment SAEs occurring after ICF signoff and before the first dosing of investigational product (OCA or placebo) will be tabulated in the same manner as above for all subjects participating in the washout period.

The following listings will be presented by subject:

- All AEs
- Serious AEs (subset of the AEs where serious is marked as "Yes")
- Death information will be provided in a separate listing, should any deaths occur
- Severe AEs (subset of AEs where severity is marked as "Severe" or severity is missing)
- Related AEs (subset of AEs where relationship to study medication is marked as "Definite", "Possible" or "Probable")
- AE's leading to withdrawal of investigational product (subset of AEs where action taken with study medication is marked as "Drug Withdrawn")
- AE's leading to Study Discontinuation (subset of AEs where subject discontinued from study is checked)

10.3. Adverse Events of Special Interest

Adverse events of special interest (AESI) are pruritus, and dyslipidaemia. For each of these AESI, subject incidence of TEAEs and the total number of entries by MedDRA system organ class and preferred term will be presented. In addition, for each AESI, the time to first onset and time to most severe event will be summarized using descriptive statistics.

10.3.1. Pruritus

Treatment-emergent pruritus, defined as any preferred term including "Prur," will be summarized separately by the MedDRA system organ class (SOC), treatment group, and preferred term (PT) as a subset of all TEAEs.

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

In 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 report an AE of pruritus will be censored at the date of last contact.

- Time to onset of the first severe treatment-emergent pruritus
  - The time to the start of the first serious pruritus will be calculated by date of onset of the first severe pruritus – date of first dose of investigational product + 1.
  - Subjects who never report a severe AE of pruritus will be censored at the date of completion or discontinuation.
- Time to resolution of the most severe treatment-emergent pruritus event
  - The time to resolution of the most severe event of treatment-emergent pruritus will be calculated by date of resolution of the most severe event of treatment-emergent pruritus – start date of the most severe event of treatment-emergent pruritus + 1.
  - Subjects whose most severe event of treatment-emergent pruritus is ongoing will be censored at the date of completion or discontinuation. For subjects that are lost to follow-up, the last contact date will be used.
  - Only subjects that report a pruritus event are included in this analysis.
  - Separate summaries will be presented for those who discontinue due to pruritus.

The analysis of time to event will include the number of subjects with the event (first onset, first moderate or severe, first severe), the number of subjects without the event (censored), descriptive statistics of time for those with an event, and range in days for all subjects. 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, with the number at risk identified. The comparison of OCA to placebo will be summarized using a log-rank test stratified by randomization strata factor. The analysis will be repeated on the subset of subjects who experienced a treatment-emergent pruritus event.

10.3.2. Dyslipidaemia

AE lipid profile changes, defined in the Dyslipidaemia SMQ, will be reported. See Appendix B for a list of the preferred terms and their associated codes. These events will be summarized separately by the MedDRA system organ class (SOC), treatment group, and preferred term (PT) as a subset of all TEAEs.

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

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

The analysis of time to event will include the number of subjects with the event (first onset) and the number of subjects without the event (censored), descriptive statistics of time for those with

an event, and range in days for all subjects. 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, with the number at risk identified. The comparison of OCA to placebo will be summarized using a log-rank test stratified by randomization strata factor.

10.4. Clinical Laboratory Evaluations

Clinical laboratory evaluations during the DB phase are assessed at the central laboratory. A listing of available laboratory reference/normal ranges for each laboratory parameter will be provided including age, sex, values with units. All analyses of lipoprotein will include fasted samples only.

Quantitative hematology, coagulation, serum chemistry, apolipoprotein and nuclear magnetic resonance (NMR) lipoprotein panel laboratory parameters will be summarized by treatment group in SI units using arithmetic summary statistics as described in Section 6.2.1 at Baseline and at each scheduled DB post-Baseline visit. Change from Baseline will also be summarized in the same manner. Re-tests or unscheduled visit results after first dose of investigational product in the DB phase will not be summarized but will be included in the data listings.

In addition, shift tables (ie, low-normal-high at Baseline versus low-normal-high at DB post-Baseline Visit in a 3-by-3 contingency table) from Baseline to worst value, last value, and at each scheduled DB post-Baseline visit will be provided for hematology, coagulation, and serum chemistry by treatment group.

Urinalysis results will not be summarized but will be provided in a data listing.

Listings of all laboratory values will flag values outside of the normal range as high (H) or low (L) and indicate whether or not a value is clinically significant (CS), based on investigator judgment.

For laboratory test results that are below the quantifiable limits:

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

For laboratory test results that are above the quantifiable limits:

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

10.5. Vital Signs

Vital signs (oral temperature, sitting heart rate, respiratory rate, and sitting blood pressure [systolic and diastolic]) will be summarized using arithmetic summary statistics as described in Section 6.2.1 at Baseline and at each DB post-Baseline visit. Change from Baseline will also be summarized in the same manner. Re-tests or unscheduled visit results prior to the administration time of the first dose in the DB phase will be included in the calculation of Baseline. Re-tests or unscheduled visit results after the first dose of investigational product administration will not be summarized but will be included in the data listings.

10.6. Electrocardiograms (ECGs)

Overall interpretation results for ECGs and the Investigator interpretation results are collected as normal, abnormal-not clinically significant, and abnormal-clinically significant. Subjects whose interpretation shifts from normal at Baseline to abnormal at any DB post-Baseline visit will be summarized by treatment group and visit. In addition, these will be listed separately including description of the abnormality and any associated comments.

11. CHANGES TO PROTOCOL-SPECIFIED ANALYSES AND ENDPOINTS

Gut microbiome and genetic samples will potentially be analyzed. Details of this analysis will be included in a separate SAP.

12. REFERENCES

D'Agostino R.B., Belanger A, D'Agostino R.B. Jr. A suggestion for using powerful and informative tests of normality. The American Statistician; 1990;44(4):316-21.

Elman S, Hynan L, Gabriel V, et al. The 5-D itch scale: a new measure of pruritus. British Journal of Dermatology. 2010;162:587-593.

Glaser, R.E. Levene's robust test of homogeneity of variances. Encyclopedia of Statistical Sciences. New York: John Wiley and Sons; 1982:608-610.