STATISTICAL ANALYSIS PLAN

A Phase 2 Open Label Proof-of-Concept Study of HTD1801 in Adult Subjects with Primary Biliary Cholangitis (PBC) and an Inadequate Response to Standard Therapy

(PRONTO-PBC)

Protocol HTD1801.PCT013
IND 136137

Sponsored by:
HighTide Biopharma Pty. Ltd.
Level 13, Citigroup Tower
2 Park Street
Sydney, NSW 2000
Australia

Version: 1.0
Date Approved: 25 Apr 2022

Confidentiality Statement:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

AASLD	American Association for the Study of Liver Diseases
AE	adverse event
AIH	autoimmune hepatitis
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AMA	antimitochondrial antibody
ANA	antinuclear antibody
AST	aspartate aminotransferase
BBR	berberine
BID	twice daily
BMI	body mass index
CBC	complete blood count
CFR	Code of Federal Regulations
CRP	C-reactive protein
ECG	electrocardiogram
eCRF	electronic case report form
EDC	electronic data capture
ELF	enhanced liver fibrosis
FSH	follicle stimulating hormone
GGT	gamma-glutamyl transferase
GI	gastrointestinal
HDL	high-density lipoprotein cholesterol
ICF	informed consent form
ICH	International Conference on Harmonisation
IND	Investigational New Drug
INR	international normalized ratio
IRB	Institutional Review Board
ITT	intention-to-treat
LDL	low-density lipoprotein cholesterol
NASH	nonalcoholic steatohepatitis
OCA	obeticholic acid
PBC	primary biliary cholangitis
PSC	primary sclerosing cholangitis
SAE	serious adverse event
SD	standard deviation
TEAE	treatment-emergent adverse event
UDCA	ursodeoxycholic acid
ULN	upper limit of normal
VAS	visual analogue scale

1 STUDY DESCRIPTION

1.1 Introduction

This document outlines the statistical methods to be implemented during the analyses of protocol HTD1801.PCT013. 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).

In this study, HTD1801 is being studied for the treatment of Primary Biliary Cholangitis (PBC) in adult subjects. PBC is a relatively rare chronic, slowly progressive, autoimmune liver disease. It can lead to inflammation, destruction of small bile ducts, fibrosis, and eventually cirrhosis. The most common symptoms of PBC are fatigue and itching.

HTD1801 is a novel synthetic small molecule: an ionic salt formed between a base [berberine (BBR)] and a weak acid [ursodeoxycholic acid (UDCA)] with a stoichiometry of 1:1 and containing a single active ingredient in the final dosage form. The single active ingredient in HTD1801 likely dissociates into two active moieties, BBR and UDCA. These moieties work in tandem in the salt form to produce, through their interaction, distinct and improved properties of HTD1801 that are not observed with either of the individual active moieties. A substantial portion of the beneficial effect of the novel HTD1801 salt appears to be derived from the interaction of its BBR and UDCA moieties conferring improved physico-chemical properties, solubility, and PK properties.

UDCA has been the mainstay of treatment for PBC for many years and is used globally as the first-line therapy for the disease. However, approximately 40 to 50% of PBC patients do not respond to UDCA or do not maintain an adequate response to UDCA.

The addition of BBR to UDCA, as a salt, in HTD1801 could deliver a benefit to PBC patients who are inadequate responders to UDCA alone. While BBR alone has not been studied in PBC, in a rat model of cholestasis and liver inflammation, HTD1801 was effective in decreasing liver inflammation and necrosis, and improving bile duct health by reducing inflammation in the portal area and controlling bile duct proliferation. In addition, treatment with HTD1801 was found to be associated with an improvement in serum markers of cholestasis (alkaline phosphatase [ALP] and gamma-glutamyl transferase [GGT]).

1.2 Objectives

The primary objective of this study is to evaluate the effects of HTD1801 on serum ALP in adult subjects with PBC who have experienced an inadequate response to standard therapy. Inadequate response is defined as ALP ≥1.5x ULN despite having been on adequate doses of UDCA for at least 6 months.

The secondary objectives of this study are to evaluate the following:

- The effects of HTD1801 on serum markers of cholestasis, serum lipids and serum markers of inflammation
- The safety and tolerability of HTD1801 over 12 weeks of treatment

- The pharmacokinetic (PK) profile of HTD1801 in a subset of subjects
    - PK analyses will be covered under a separate analysis plan.

1.3 Study Design

This is a Phase 2, single arm, 12-week open-label proof of concept study of HTD1801 in adult subjects with PBC with an inadequate response to standard therapy. Inadequate response is defined as ALP ≥1.5x ULN despite having been on adequate doses of UDCA for at least 6 months.

The study will consist of three periods, for a total study period of 16 weeks:

- Screening period: Up to 28 days
- Treatment period: 12 weeks
- Follow-up period: 28 days after the last dose of study drug

Upon confirmation of eligibility, approximately 30 subjects with PBC will be enrolled and will receive HTD1801 1000 mg twice daily (BID) for 12 weeks. Study visits occur at Day 0, Weeks 2, 4, 6, 8, and 12. A final follow-up visit will occur at Week 16, 4 weeks after the last dose of study drug. If a subject experiences gastrointestinal discomfort, as outlined in the protocol, at the starting dose of 1000 mg BID, dose reduction is allowed.

An abbreviated Schedule of Events for the study can be found in Table 1.

.

Table 1 Schedule of Events

Examination	Screening	Baseline	Week 2 Clinic/ Optional Home Visit	Week 4 Clinic/Optional Home Visit	Week 6	Week 8 Clinic/Optional Home Visit	Week 12	ET	Week 16 F/U
Demographics	X								
Pruritus VAS	X	X			X		X	X	
GLOBE Score		X			X		X	X	
ECG	X						X	X	
Vital signs	X	X			X		X	X	X
Medical history	X	X							
Physical exam	X	X			X		X	X	X
Prior and Concomitant medications	X	X	X	X	X	X	X	X	X
Hematology	X	X			X		X	X	
Serum Chemistry panel (including GGT)	X	X	X	X	X	X	X	X	X
Endocrine/Reproductive	X								
Serology, Other labs	X	X			X		X	X	
Lipid panel	X	X			X		X	X	
Bile acid panel[a]		X			X		X	X	
PK Blood Draw					X		X		
Urine pregnancy test		X					X	X	
Dispense / collect study drug		X			X		X	X	
AEs/SAEs		X	X	X	X	X	X	X	X
[a]For subjects that had a sample obtained at Baseline									

1.4 Treatment Blinding

This is an open-label study.

1.5 Decision Rule and Sample Size

The significance level for this study is 5% (one-sided). If the one-sided p-value is less than 5% then the null hypothesis of no treatment effect is rejected.

A sample size of 30 subjects provides ≥90% power for the primary endpoint assuming a mean percent reduction in ALP of 15% and a standard deviation of 25%. Power calculations use a 5% one sided alpha and a t-test. Sample size calculation were performed in PASS.

At least 6 subjects will participate in a PK sub study.

2 STATISTICAL METHODS

2.1 Analysis Populations

The following analysis populations will be defined for statistical analysis:

- Screened population - All subjects who sign the study informed consent.
- Intention-to-treat (ITT) population - All subjects that receive at least one dose of HTD1801. This population will be used for the primary endpoint, as well as all efficacy, safety, and baseline/demographic analyses and summaries.
- Per Protocol (PP) Population – All subjects in the ITT population who provide a Week 12 assessment and 80% study drug compliance. This population will be used for selected efficacy analyses and summaries.

2.2 Study Drug Dosing and Compliance

HTD1801 tablets are administered BID with water and food. Subjects are instructed to take the entire contents of 1 pouch each morning and each evening with water and food (e.g., meal or snack). An adequate supply of pouches (containing 4 identical white tablets [250 mg per tablet]) will be placed into kits packaged for weekly use.

Study drug will be dispensed at Baseline and the Week 6 visit. Each subject will receive the appropriate number of kits to provide enough supplies until the subsequent visit where study drug is dispensed. At the Week 6 and Week 12 visits subjects will return any unused study drug to the investigational site. At these visits, the number of full and empty or partially empty pouches will be captured on the eCRF.

The tablet counts reported within the eCRF will be used as the dosing and compliance data source. From these data the duration of treatment, number of tablets used, average dose received per day, and compliance with treatment will be determined. The duration of treatment is calculated as,

$$\textit{Treatment Duration} = \textit{treatment end date} - \textit{treatment start date} + 1.$$

Compliance will be calculated as,

$$\text{CCCCCCCCCCCCCCCCCCCC(\%)} = \frac{\text{TTCCTTCCCCTTTT UUTTCCUU}}{\text{TTCCTTCCCCTTTT EEEECCCCCCTTCCUU TTCC TTCC UUTTCCUU}} .$$

The number of tablets used will be based upon the counts recorded in the eCRF. The total number of tablets used will be the sum across kits. Should study drug reconciliation data be missing for a subject, the number of tablets used will also be missing, and therefore compliance will be unknown.

If there is no dose interruption or modification, tablets expected to be used is,

$$\textit{Tablets Used} = 8 * (\textit{treatment end date} - \textit{treatment start date} + 1)$$

If there was a period of time where there was a dose interruption or modification, the dose will be assumed to be 0 for an interruption, and the modified dose will be the number of tablets received per day of modification.

This compliance calculation implies a subject who takes treatment for one week but takes all expected tablets for that week will be 100% compliant with treatment.

Average dose received per day will be calculated as,

$$AAAACCAACCAACC\ DDCCTTCC\ RRCCCCCCCCAACCUU\ CCCCAA\ DDCCDD = \frac{TTCCTTCCCCTTTT\ UUTTCCUU}{TTAACCCCTTCCCCCCTT\ DDDDAACCTTCCCCCC - DDDDAACCTTCCCCCC\ CCoo\ IICCTTCCAAAADDCCTTCCCCCC}.$$

2.3 Study Endpoints

The study endpoints are listed below, with greater detail concerning these endpoints provided in Sections 2.3.1 to 2.3.8.

Primary Efficacy Endpoint

The primary efficacy endpoint is the percent change in ALP at Week 12 compared to Baseline.

ALP Secondary Efficacy Endpoints

The ALP secondary efficacy endpoints are as follows:

- Absolute change in serum ALP at Week 12 compared to Baseline
- Proportion of subjects who have ≥20% decrease in ALP from Baseline to Week 12
- Proportion of subjects who have ≥40% decrease in ALP from Baseline to Week 12
- Proportion of subjects who normalize ALP at Week 12

Secondary Endpoints

- Change in total bilirubin from Baseline to Week 12
- Change in serum GGT between Baseline and Week 12
- Change in alanine aminotransferase (ALT) from Baseline to Week 12
- Change in aspartate aminotransferase (AST) from Baseline to Week 12
- Change in total cholesterol, low-density lipoprotein (LDL) cholesterol, and triglyceride levels between Baseline and Week 12
- Change in fibrinogen, c-reactive protein (CRP), haptoglobin, enhanced liver fibrosis (ELF) score, and serum immunoglobulins between Baseline and Week 12

Safety endpoints include the following:

- Changes from Baseline to Week 6 and Week 12 in GLOBE Score

- Changes from Baseline to Week 6 and Week 12 in pruritus (average and worst) as measured by pruritus visual analog score (VAS)
- Changes from baseline in vital signs, and clinical laboratory values
- Incidence of treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs)
- PK parameters including $T_{1/2}$, C_{max}, and AUC
    - The analysis of these PK endpoints will be performed outside of this analysis plan.

Additional Analyses

- Proportion of subjects who have ≥10% decrease in ALP from Baseline to Week 12
- Proportion of subjects at Week 12 who had Baseline ALP ≥1.67x ULN and have ALP <1.67x ULN, ≥15% decrease in ALP and total bilirubin ≤ULN at Week 12.
- A plot of percent change in ALP at each timepoint through Week 12 for the ratio of current to prior dose of UDCA
- Proportion of subjects that crossed the liver monitoring thresholds will be summarized. The liver monitoring threshold is defined as ALT:
    - 3x ULN if the Baseline ALT is ≤60 U/L, or
    - 3x Baseline ALT if Baseline ALT is >60 U/L

Additional analyses are not listed as endpoints in the protocol but are listed here as prespecified extensions of the protocol endpoints.

2.3.1 Alkaline Phosphatase

ALP results are provided by the central and local laboratory with units of Units per Liter (U/L). Baseline is defined as the last available ALP result prior to treatment with change from baseline and percent change from baseline based upon this baseline value as the respective timepoint. Normalized ALP is defined as having ALP within the normal range (37 U/L to 116 U/L). Historical ALP will be summarized in the baseline characteristics table but not used otherwise.

2.3.2 Liver Laboratory Tests

Total bilirubin, GGT, AST, and ALT results are provided by the central and local laboratory. The units these parameters are provided in and the normal range for these parameters at the central laboratory are provided below.

Laboratory Test	Units	Normal Range
Total bilirubin	mg/dL	0.25 to 1.21
GGT	U/L	7 to 38
AST	U/L	14 to 43
ALT	U/L	10 to 53

2.3.3 Lipid Laboratory Tests

Total cholesterol, low-density lipoprotein (LDL) cholesterol, and triglyceride results are provided by the central and local laboratory. The units these parameters are provided in and the normal range for these parameters are provided below.

Laboratory Test	Units	Normal Range
Total Cholesterol	mg/dL	100 to 200
LDL	mg/dL	50 to 130
Triglycerides	mg/dL	50 to 150

2.3.4 Serum Markers of Inflammation

Fibrinogen and c-reactive protein (CRP) results are provided by the central and local laboratory. Haptoglobin and serum immunoglobulins results are provided by the central laboratory. The units these parameters are provided in and the normal range for these parameters are provided below.

Laboratory Test	Units	Normal Range
Fibrinogen	mg/dL	200 to 400
CRP	mg/L	0 to 3
Haptoglobin	g/L	0.3 to 2
Immunoglobulin G	mg/dL	700 to 1600
Immunoglobulin M	mg/dL	40 to 230

2.3.5 Enhanced Liver Fibrosis Score

The ELF score is a set of serum markers that consist of hyaluronic acid (HA), tissue inhibitor of metalloproteinases 1 (TIMP-1) N-terminal pro-peptide of type III collagen (PIIINP), and tissue inhibitor of metalloproteinases 1 (TIMP-1) used to predict moderate fibrosis and cirrhosis (Lichtinghagen et al., 2013). Change from Baseline to Week 6 and Week 12 will be presented, along with a shift table for the categorical interpretation. The interpretation of the ELF score is as follows:

- None to mild liver fibrosis: ELF score <7.7
- Moderate liver fibrosis: ELF score ≥7.7 and <9.8
- Severe liver fibrosis: ELF score ≥9.8
- Cirrhosis: ELF score >11.3

2.3.6 GLOBE Score

The GLOBE score is a validated risk assessment tool, able to accurately stratify PBC patients to high and low risk for future adverse events (AEs), where a higher GLOBE score is considered higher risk (Lammers , 2015). Change from Baseline to Week 6 and Week 12 will be presented, along with a shift table for the categorical interpretation, as well as Survival Risk at 5, 10, and 15 years. The GLOBE score will be calculated using the following algorithm:

$$GLOBE\ Score = 0.044378 \times AAAACC + 0.93982 \times CCCC^{\frac{BBBBBBBBBBBBBBBBB}{UUUUUU}} + 0.335648 \times CCCC^{\frac{AAUUAA}{UUUUUU}} - 2.266708 \times {}^{\frac{AABBBBBBAABBBB}{UUUUUU}} = 0.002581 \times PPCCCCTTCCCCTT\ CCCCDDCCTT + 1.216865$$

Age (years): at initiation of UDCA therapy.

Units: Bilirubin (µmol/L or mg/dL), ALP (U/L), Albumin (g/L), Platelets (10^9/L)

$$SSDDAAAACCAACCCC\ \ RRCCTTRR\ \ SSCCCCAACC = 1 - BBCCTTCCCCCCCCCC\ \ SSDDAAAACCAACCCC\ \ FFDDCCCCTTCCCCCC(CC^{GGUUGGBBGG\ \ SSSSSSBBSS})$$

Baseline survival function = 0.9385 (at 5 years); 0. 8429 (10 years); 0.7361 (15 years).

The interpretation is as follows:

- Higher risk: GLOBE score >0.3
- Lower risk: GLOBE score <0.3

2.3.7 Pruritus Visual Analog Scale

The Pruritus VAS is a validated, self-reported, instrument for measurement of itch intensity. It uses a 24-hour recall period and asks subjects to rate the average and worst intensity of their itch on a 10 cm horizontal line ranging from 0 cm (no itch) to 10 cm (worst imaginable itch). Change from Baseline to Week 6 and Week 12 will be presented for subjects with a value of ≤0.5cm at baseline and >0.5cm at baseline. The proportion of subjects with improvement (for subjects >0.5cm at baseline) and worsening (for subjects ≤0.5cm at baseline) will also be presented, along with a shift table for the categorical interpretation. The interpretation for the Pruritus VAS score is as follows:

- No pruritus: VAS=0
- Mild pruritus: VAS <3
- Moderate pruritus: VAS 3 to <7
- Severe pruritus: VAS 7 to <9
- Very severe pruritus: VAS ≥ 9

2.3.8 Safety Parameters

The following safety parameters are captured within this study.

2.3.8.1 Adverse Events and Serious Adverse Events

AEs and SAEs will be recorded starting after Screening and at every visit through the Week 16 Follow-Up Visit. An AE is considered treatment emergent, TEAE, if it began after study drug administration.

AEs will be graded for severity using the Common Terminology Criteria for Adverse Events (CTCAE) v5.0 to describe the maximum intensity of the AE. As such, severity will be reported as Mild, Moderate, Severe, Very Severe, Life Threatening or Disabling, Death Related. Causality of an AE will be assessed by the Investigator using the following terms: Possibly Related and Unrelated.

AEs and medical history will be coded using MedDRA version 24.0.

2.3.8.2 Adverse Events of Special Interest

Adverse events of special interest (AESI) include gastrointestinal events, treatment failure due to dose reduction of study drug as defined in the protocol.

2.3.8.3 Other Safety Endpoints

- Vital signs, height, and weight:
    - Blood pressure, pulse, respiratory rate, height, weight, body temperature, and the presence of clinically significant abnormal values determined by an Investigator. Clinically significant abnormalities identified after the start of treatment will be reported as AEs.
- 12-Lead electrocardiogram (ECG):
    - An Investigator assessment of normal vs abnormal will be available as well as clinically significant vs not clinically significant. Abnormal, clinically significant ECG readings seen after the start of treatment will be recorded as AEs.
- Laboratory samples will include:
    - Hematology, serum chemistry, lipids, coagulation, and other. Clinically significant changes in laboratory values will be recorded by the clinical sites as AEs.
- Post-Treatment Withdrawal Flares:
    - The difference between Week 12 and Week 16 laboratory values for ALP, ALT, AST, and GGT will be summarized.

2.4 Study Day and Visit Windows

Summary tables will report data based upon the protocol scheduled time points (Baseline, Weeks 2, 4, 6, 8, 12, 16). Assessment will be assigned to these time points based upon the study day they are performed on. Study day is defined as

Study Day = date of assessment - date of first treatment.

The analysis windows used to report endpoints are outlined in Table 2. Assessments are assigned to an analysis window based upon the study day it occurred on as well as the treatment status at the time the assessment was performed. If more than one assessment is available within the range, the assessment closest to the target day is reported for the analysis window. If two observations exist with the same distance to the target day, the first observation is used.

Table 2 Analysis Windows

Visit	Study Day Range	Target Day	Treatment Status
Screening	< Day –1	First	Off Treatment[a]
Baseline Visit	Day 1	Day 1	Off Treatment[a]
Week 2	Day 2 to Day 21	Day 14	On Treatment[b]
Week 4	Day 22 to Day 35	Day 28	On Treatment[b]
Week 6	Day 36 to Day 49	Day 42	On Treatment[b]
Week 8	Day 50 to Day 70	Day 56	On Treatment[b]
Week 12	Day 71 to Day 90	Day 84	On Treatment[b]
Week 16	> Day 90	Day 112	Off Treatment[c]

[a] Prior to first treatment.
[b] After the start of treatment to 1 day after the last dose.
[c] More than 1 day after the last dose.

2.4.1 Definition of Baseline

Unless specified elsewhere, baseline is defined as the last available measurement prior to administration of study drug. For events that occur on the same day as the first administration of study drug (i.e. assessments performed at the Baseline visit), and for which the time of the assessment is not available, the assessment will be classified as pre-treatment, hence baseline. The exception to this rule will be AEs for which the event will be classified as a TEAE.

It should be noted that for the purpose of clinical management, the description of baseline varies in the protocol (e.g. inclusion/exclusion and Drug-Induced Liver Injury Monitoring). For analysis purposes, only the definition above will be used.

2.5 Statistical Assessment of the Study Objectives

The efficacy endpoints will be summarized within the ITT population using descriptive statistics, with the primary endpoint, GGT and total bilirubin, also summarized within the PP

population. As this study includes one treatment group, the hypotheses, tests, and confidence intervals are all one sample comparisons to baseline. Confidence intervals will be 95% two-sided normal approximation confidence intervals. Testing for the percent change from Baseline to Week 12 in ALP using a t-test will also be performed. Testing will be one sided using a 5% alpha level. No multiplicity adjustment will be used, and the p-values reported for the endpoints will be descriptive in nature.

2.5.1 Primary Analysis

2.5.1.1 Hypothesis and Estimand

The primary hypothesis for this study is that treatment with HTD1801 results in reductions in ALP:

$$HH_0: \frac{AAAAPP_{WWSSSSWW\ 12} - AAAAPP_{BBBBBBSSBBBBBBSS}}{AAAAPP_{BBBBBBSSBBBBBBSS}} \geq 0\ AATT.\ HH_{BB}: \frac{AAAAPP_{WWSSSSWW\ 12} - AAAAPP_{BBBBBBSSBBBBBBSS}}{AAAAPP_{BBBBBBSSBBBBBBSS}} < 0$$

The estimand will have the following characteristics:

- The target population will be made up of all subjects on treatment who provide data at Week 12
- The endpoint will be the percent change from baseline in ALP
- Summarized results are limited to on treatment results at Week 12 (i.e., results from prior to Week 12 are not included in the analysis of the Week 12 results). Hence intercurrent events that take subjects off treatment or prevent ALP data from being available at Week 12 will remove subjects from the analysis.
- The endpoint will be summarized as the mean percent change from baseline in ALP

2.5.1.2 Analysis Methods

Serum ALP levels, change from baseline, and percent change from baseline will be summarized at each analysis window (see Table 2) using descriptive statistics and 95% confidence intervals. A t-test will also be performed to test for the percent change from baseline to Week 12. This analysis will be based on observed data without imputation. This analysis will be summarized within the ITT and PP population.

2.5.2 Secondary Analyses

Secondary endpoints, including the secondary ALP endpoints, will be summarized with descriptive statistics by timepoint. These summary measures will include baseline, result at the time point that data is captured, change from baseline, and percent change from baseline as is appropriate for the endpoint. A 95% confidence interval will be produced for changes in laboratory endpoints, within the ITT population.

Endpoints with categorical interpretations (e.g. ELF, GLOBE Score, Pruritus VAS) will additionally be summarized by shift tables for Change from Baseline to Week 12 using the categories listed above.

2.5.3 Subgroups

The results, change from baseline, and percent change from baseline to week 12 for ALP, total bilirubin, GGT, ALT, AST, and ELF will be performed for the following subgroups:

- ALP at Baseline (above and below median value)
- Total bilirubin at Baseline (above and below median value)

2.6 Handling of Missing Data

Summary statistics will generally be reported based upon observed data. Missing data will not be imputed for the efficacy analyses. Percentages will be reported for subjects with data. Missing dates for AEs and concomitant medications will be imputed to allow these items to be summarized as pre-treatment or not. Should a determination of treatment period (on treatment or pre-treatment) be required for AEs or concomitant medication, but the corresponding date is missing or partial, the event/medication will be considered on treatment unless the portions of the date that are available indicate this is not possible.

2.7 Safety Analyses

Safety endpoints will be summarized with descriptive statistics. All safety summaries and analyses will be performed using the ITT population.

Prior and concomitant medications and procedures will be coded by the World Health Organization (WHO) Drug Dictionary version Sept2020. Medications and procedures will be considered concomitant unless they ended prior to treatment.

2.8 Interim Analyses

Data will be summarized at study completion, as well as when 15 subjects have reached Week 12. At the interim timepoint, the data will be summarized descriptively: primary efficacy endpoint, additional ALP secondary endpoints, and TEAEs. For the interim analysis, only subjects with completed data who reached Week 12 will be used for efficacy analyses, all available data will be used for safety analyses.

Enrollment into the study may be stopped early prior to enrollment of all 30 subjects should the scientific goal (evaluate the effects of HTD1801 on serum ALP) of the study be achieved with the currently available data. That is, if prior to enrollment of 30 subjects, the results provide sufficient preliminary evidence of efficacy that warrant further evaluation in a placebo-controlled study, enrollment in this study may be terminated.

3 SUMMARY TABLES, LISTINGS, AND FIGURES

3.1 General Conventions

Unless otherwise stated, summary statistics including the number of subjects (n), mean, standard deviation, median, minimum, and maximum will be presented for continuous variables. Generally, the minimum and maximum values will be presented to the same decimal precision as the raw values, the mean and median values to one more, and the standard deviation, to two more decimal places than the raw values. For categorical variables, per category, the absolute counts (n) and percentages (%) of subjects with data will be presented. Percentages will be presented to one decimal place.

For AEs, medical history and concomitant medications will be reported on a per-subject basis. The denominator for the percentage calculation will be the number of subjects at risk in each treatment group. A subject will be considered at risk if the subject is in the analysis set and in the subgroup of interest.

All p-values will be rounded to 4 decimal places. If a p-value is less than 0.0001 it will be displayed as "< 0.0001", a p-value rounding to 1 will be displayed as "> 0.9999".

3.2 Clinical Study Subjects

3.2.1 Subject Disposition and Analysis Populations.

The number of subjects treated and discontinued early (i.e., follow-up visit was not completed), will be summarized. The reason for early termination will be summarized. Additionally, a summary of subjects who attended each visit result will be produced.

A listing of entry criteria that were not met will be produced. The listed criteria will include the criteria language. A table of the violated entry criteria for subjects will be produced, as well as a table that summarizes the screen failure reasons for screen failure subjects.

The number of subjects in each analysis population will be summarized.

3.2.2 Demographics and Baseline Characteristics

Demographic and baseline characteristics will be summarized descriptively. This will include the following items:

- Age at informed consent
- Race
- Ethnicity
- Sex
- Height
- Weight

- BMI
- Duration of Primary Biliary Cholangitis
- Biochemical evidence of Cholestasis (Yes or No)
- Age of initiation of UDCA
- Age at diagnosis of PBC
- Most recent UDCA dose (mg/kg/day)
- Presence of antimitochondrial antibody
- Historical ALP
- Pruritus VAS

3.2.2.1 Baseline Laboratory Assessments

Baseline laboratory assessments will be summarized descriptively. This will include:

- ALP
- GGT
- Total Bilirubin
- AST
- ALT
- Albumin
- Platelets
- AMA (Positive, Equivocal, or Negative)
    - For positive values, AMA will also be summarized as a continuous measure
- ANA (Positive, Equivocal, or Negative)
    - For positive values, ANA will also be summarized as a continuous measure

3.2.3 Concomitant Medications

Medications with a stop date before the treatment dosing date will be considered prior medications/procedures. Medications with a start or stop date on or after the treatment dosing date will be considered concomitant. All medications marked as ongoing are concomitant.

A medication with an incomplete stop date will be considered concomitant if:

- Month is missing and year is equal to or after the year of treatment dosing date
- Day is missing and year is equal to the year of the treatment dosing date and month is equal to or after the month of the treatment dosing date.

All concomitant medications will be summarized by preferred terms in descending order of frequency. All medications, including prior and prohibited medications, will be provided in a listing.

3.2.4 Medical History

Medical history will be tabulated by system organ class and preferred terms and provided in a listing.

3.2.5 Clinical Study Treatment and Extent of Exposure

The total tablets used, duration of treatment, and compliance (see Section 2.2) will be summarized. Average dose received per day, along with the number of subjects with dose interruption or modification will also be summarized.

3.3 Analysis of Efficacy Endpoints

3.3.1 Serum Alkaline Phosphatase

The serum ALP levels, change from baseline, and percent change from baseline will be summarized with descriptive statistics by timepoint for the ITT and PP populations. The summary measures will include 95% normal distribution confidence intervals of the treatment difference. In addition, a t-test will be performed for the percent change from baseline to Week 12. Subgroup analyses (see Section 2.5.1.2) will be performed for the change from baseline and percent change at Week 12 data.

The proportion of subjects who have ≥20% and the proportion with a ≥40% decrease in ALP from Baseline will be summarized at each timepoint with Week 12 being the study specified timepoint. Additionally, the proportion of subjects who normalize ALP will be summarized by timepoint.

3.3.2 Secondary Endpoints

Tables for the secondary endpoints, see below, will be summarized for the respective laboratory values and the change from baseline results at each timepoint. Total bilirubin and GGT will be summarized within the ITT and PP populations. Subgroup analyses for specified laboratory parameters (see section 2.5.2) will be performed for the change from baseline and percent change at Week 12. Central and local lab data will be used for the efficacy analyses.

- Total bilirubin
- GGT
- ALT
- AST
- Total cholesterol

- LDL
- Triglyceride
- Fibrinogen
- CRP
- Haptoglobin
- ELF
- Immunoglobulins (IgM and IgG)

3.3.3 UDCA Use

A plot of average percent change in ALP at each timepoint through Week 12 for groups defined by the ratio of current to prior dose of UDCA. The ratio groups displayed in the plot will be <90%, 90% to 110%, and >110%. Separate subject level plots will also be presented for each ratio group through Week 12.

3.3.4 ELF Score

The number of subjects within ELF score interpretations will be presented at Week 12 as well as the shift within categories for change from baseline to week 12. Summary statistics at Week 12 will also be presented.

3.3.5 GLOBE Score

The number of subjects within GLOBE score interpretations will be presented along with the Fibrosis Score interpretations. Change in GLOBE score from baseline will be summarized using descriptive statistics at Week 6 and Week 12. A listing of the GLOBE score will also be produced and summarized by duration of UDCA exposure.

Survival risk (%) for 5, 10, and 15 years will be summarized using descriptive statistics at Week 12 for the GLOBE Scores.

3.4 Analysis of Safety Endpoints

Safety Endpoints (Adverse Events, Clinical Laboratory Results, Vital Signs) will be presented at the scheduled timepoints as outlined in Table 1, with further detail in the protocol.

3.4.1 Adverse Events

An overview of AEs, which includes subject incidence of TEAEs, treatment-related TEAEs, serious TEAEs, AESIs, TEAEs by CTCAE grade and TEAEs leading to permanent discontinuation of treatment. For TEAEs presented by CTCAE grade, the worst grade during the clinical study will be presented for each subject.

The subject incidence of TEAEs, treatment-related TEAEs and TEAEs leading to permanent discontinuation of treatment will be summarized by system organ class and preferred term.

TEAEs will also be summarized in a table by CTCAE grade. For TEAEs presented by CTCAE grade, the worst grade for each event during the clinical study will be presented for each subject. A listing for all off-treatment AEs will be presented similarly.

All AEs will be presented as a listing by subject. TEAEs leading to treatment discontinuation will be provided in a separate listing.

3.4.1.1 Post-Treatment Withdrawal Flares

The difference between Week 12 and Week 16 laboratory values for ALP, ALT, AST, and GGT will be summarized using descriptive statistics.

3.4.2 Pruritus VAS

Change in Pruritus VAS will be summarized using descriptive statistics for the result and change from baseline. The proportion of subjects with improvement will be presented for subjects >0.5cm at baseline. The proportion of subjects with worsening will be presented for subjects ≤0.5cm at baseline.

3.4.3 Clinical Laboratory Evaluations

Clinical laboratory results, including bile acids, will be summarized by type of laboratory assessment (hematology, serum chemistry, lipids, coagulation, other) and timepoint. Both central and local laboratory results will be used for the safety lab summaries. Summary statistics for actual values and changes from baseline will be tabulated by analysis window. Out of range clinical laboratory values based upon the normal range will be tabulated for each timepoint. These tables will utilize the normal ranges provided for the individual sample.

The proportion of subjects that cross the liver monitoring thresholds will be summarized. A listing of ALT values for subjects that meet the liver monitoring threshold will be presented similarly to the clinical laboratory listing. Baseline values will be summarized using last available measurement prior to administration of study drug.

3.4.4 Vital Signs

The observed data at baseline and change from baseline for each measurement day will be summarized with descriptive statistics.

3.4.5 Electrocardiogram Results

The overall ECG assessment (abnormal or normal) will be summarized by analysis window. The QT values will be summarized by analysis window as both the QT value and change from baseline.

3.4.6 Pregnancy

A listing of positive pregnancy tests results will be produced.

4 REFERENCES

Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and validation of a scoring system to predict outcomes of patients with primary biliary cirrhosis receiving ursodeoxycholic acid therapy. Gastroenterology. 2015;149(7):1804-1812.e4.

Lichtinghagen R, Pietsch D, Bantel H, et al. The Enhanced Liver Fibrosis (ELF) score: normal values, influence factors and proposed cut-off values. J Hepatol. 2013;59(2):236-42.

5 APPROVAL SHEET

Product:	HTD1801
Protocol Number:	HTD1801.PCT013
SAP Version:	1.0
Version Date:	4/25/2022

The individuals signing below have reviewed and approve this statistical analysis plan.

Abigail Flyer, M.S.
Project Statistician — Date: 04/25/2022

¾eillrman, Ph.D
Statistician — Date: 04/25/2022

DocuSigned by:
Cathryn Bennett
F735F155A5B6407...

Cathryn Bennett, B.N., R.N., C.C.R.A.
VP, Clinical Operations — Date: 4/25/2022

DocuSigned by:
Adrian Michael Di Bisceglie
7F511879616A492...

Adrian M. Di Bisceglie M.D., FACP, FAASLD
Chief Medical Officer — Date: 4/25/2022