Official Title: A Phase 2, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of ACP-044 for Acute Postoperative Pain Following Orthopedic Surgery (Bunionectomy)

NCT Number: NCT04855240

Document Date: 14 March 2022

STATISTICAL ANALYSIS PLAN

Protocol No.:	ACP-044-004
Protocol Title:	A Phase 2, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of ACP-044 for Acute Postoperative Pain Following Orthopedic Surgery (Bunionectomy)
Drug:	ACP-044
Sponsor:	Acadia Pharmaceuticals Inc. [redacted]
Version No. and Date	Version 1.0, 14 March 2022

SIGNATURE/APPROVAL PAGE

AUTHOR

See appended electronic signature page _____

[redacted] Date
Senior Director, Biostatistics
Acadia Pharmaceuticals Inc.

APPROVERS

See appended electronic signature page _____

[redacted] Date
Executive Director, Clinical Science
Acadia Pharmaceuticals Inc.

See appended electronic signature page _____

[redacted] Date
Executive Director, Biostatistics
Acadia Pharmaceuticals Inc.

Signature Page for ACP-044-004 Statistical Analysis Plan

Approve	[redacted]2 21:01:35 GMT+0000
Approve	[redacted]-2022 21:24:24 GMT+0000
Approve	[redacted]2 21:41:00 GMT+0000

Signature Page for ACP-044-004 Statistical Analysis Plan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AE	adverse event
ANOVA	analysis of variance
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
AUC	area under the curve
BMI	body mass index
CMH	Cochran-Mantel Haenszel
COVID-19	Coronavirus disease 2019
C-SSRS	Columbia-Suicide Severity Rating Scale
ECG	electrocardiogram
eCRF	electronic case report form
EOT	end of treatment
ET	early termination
FAAM	Foot and Ankle Ability Measure
GESD	Global Evaluation of Study Drug
GSD	Guidance for Site Documentation and Data Management Querying of Data Impacted by COVID-19
ICE	intercurrent event
MAR	Missing at random
MCMC	Markov chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MME	Morphine milligram equivalents
MNAR	Missing not at random
msec	milliseconds
NRS	numerical rating scale
PCI	potentially clinically important
ORAE	opioid-related adverse event
QTcB	QT interval corrected for heart rate using Bazett's formula
QTcF	QT interval corrected for heart rate using Fridericia's formula
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class

TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event

1 INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the statistical analyses of safety and efficacy data as described in the study protocol Version 1.0 amendment 5 dated 01 February 2022.

Specifications for tables, figures, and listings are contained in a separate document. Statistical analyses for population pharmacokinetic (PK) and PK/pharmacodynamics (PD) modeling will be presented in a separate report and therefore will not be included in this SAP.

This plan should be read in conjunction with the study protocol and the electronic case report forms (eCRF).

2 OBJECTIVES

2.1 Primary Objective

The primary objective of this study is to evaluate the efficacy of ACP-044 compared with placebo in the treatment of acute postoperative pain.

2.2 Secondary Objectives

The secondary objectives of this study are to evaluate the efficacy of ACP-044 compared with placebo in the treatment of acute postoperative pain and to evaluate opioid use among ACP-044 treated subjects compared with placebo in the treatment of acute postoperative pain.

2.3 Exploratory Objectives

The exploratory objectives of this study are to evaluate the proportion of subjects with severe pain and to compare the impact of ACP-044 on activities of daily living compared with placebo in the treatment of acute postoperative pain.

2.4 Safety Objective

The safety objective of this study is to evaluate the safety and tolerability of ACP-044 compared with placebo in the treatment of acute postoperative pain.

2.5 Pharmacokinetic Objective

The pharmacokinetic objective of this study is to characterize the pharmacokinetic (PK) profile of ACP-044 in subjects treated for acute postoperative pain.

2.6 Pharmacokinetic/Pharmacodynamic Objective

The pharmacokinetic/pharmacodynamic objective of this study is to characterize the exposure-response relationship using appropriate modelling and simulation methods.

3 STUDY DESIGN

3.1 General Study Design

This study will be conducted as a Phase 2, randomized, double-blind, placebo-controlled, multicenter study in subjects with acute postoperative pain (following bunionectomy surgery). The study will compare each of the two active treatment groups receiving a total daily dose of 1600 mg ACP-044, administered either once daily or 400 mg every 6 hours, with a placebo group. The Sponsor, subjects, and Investigators will be blinded to treatment assignment.

Approximately three sites in the United States (US) will screen approximately 375 subjects and approximately 240 subjects are planned for enrollment (80 subjects per treatment group).

The duration of participation for individual study subjects will be approximately 9 weeks, consisting of a screening period of up to 4 weeks, a 4-day treatment period, and a safety follow-up period of approximately 30 (+4) days.

The study periods are:

- Screening: up to 4 weeks (prior to surgery)
- Double-blind treatment: Day -1 (post first dose of study drug/day of surgery) to Day 4
- Safety follow-up: approximately 7 days after end of treatment (EOT)/early termination (ET) and approximately 30 days (post last dose of study drug)

The study start date is defined as the date the first subject signs the informed consent form.

The study completion date is defined as the date the final subject, across all sites, completes their final protocol-defined assessment. Please note: this includes the safety follow-up visit or contact.

Screening Period

During the Screening period, subjects will be assessed for study eligibility. Only those subjects who meet all inclusion and no exclusion criteria will be eligible for the study.

All prohibited medications should be discontinued during the Screening period and prior to surgery. Investigators must not withdraw a subject's prohibited medication for the purpose of enrolling them into the study. Medications will be discontinued only if it is deemed clinically appropriate to do so and in consultation with the prescribing physician.

With approval of the Medical Monitor, rescreening of a subject will be allowed one time if there is an extenuating circumstance that was not under the subject's control or if is due to an

identifiable cause, and it is medically appropriate to address that cause, and there is no medical reason that would prevent the subject to be rescreened and enrolled in the study.

Treatment Period

The first dose of study drug will be administered within 60 (±15) minutes prior to the start of surgery. Subjects will then have a popliteal and/or ankle nerve block administered to anesthetize the knee, distal leg, and foot, after which they will enter into surgery. Following surgery, subjects will be administered the remaining three daily doses approximately every 6 hours after the initial dose.

Day 1 is defined as the day after surgery. The popliteal and/or ankle nerve block will be discontinued approximately 2-4 hours prior to the first dose on Day 1 (i.e., the fifth dose of the study) and nerve block removal; a discontinuation window of approximately 2 hours is permitted. The time at which the first dose on Day 1 is administered will be time 0 for all pain assessments. Study drug administration and study procedures are to be completed on Days -1 through 4 according to the schedule of assessments (Section 21.1). Subjects will be encouraged to refrain from taking rescue medication until 1-2 hours following the first dose of study drug on Day 1.

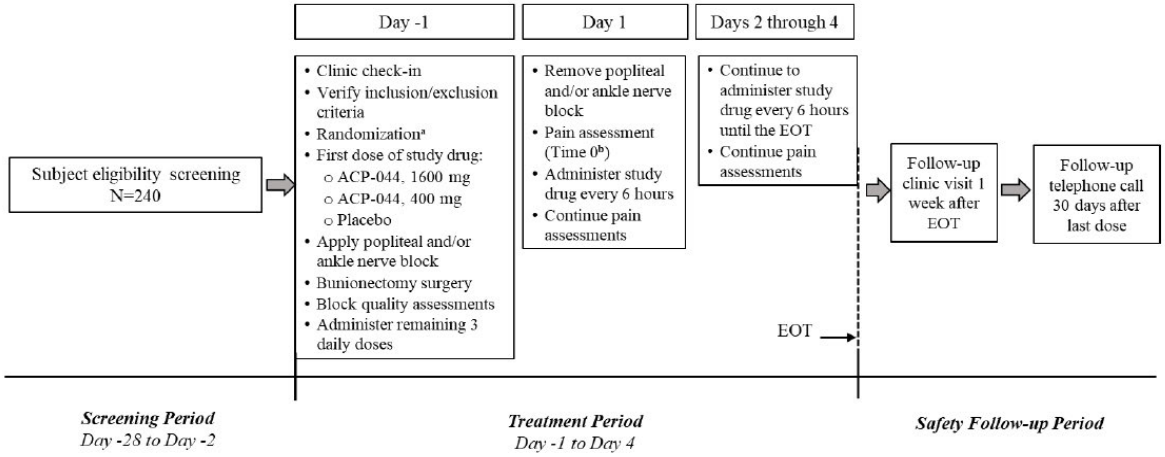
On Day 4, after the 72 hour pain assessments are performed, the EOT procedures will be completed and subjects will be discharged from the clinical site.

Safety Follow-up Period

Safety will be further assessed during a follow-up visit to occur in the clinic approximately 7 (±3) days after EOT/ET procedures and subjects will receive a follow-up telephone call approximately 30 (+4) days after the last dose of study drug. Subjects should return to standard of care after the follow-up visit that will occur in the clinic 7 (±3) days after the EOT/ET.

The study schematic is presented in Figure 1.

Figure 1 Schematic of Study Design



Abbreviations: EOT=end of treatment

[a] Subjects to be randomized in a 1:1:1 ratio to receive either a once-daily oral dose of 1600 mg ACP-044, four oral doses of 400 mg ACP-044, or placebo. Randomization will occur just prior to surgery when the first dose of study drug is given.

[b] The time at which the first dose on Day 1 (i.e., the fifth dose of study drug) is administered will be considered time 0 for pain assessments.

3.2 Schedule of Assessments

Schedule of events and assessments can be found in Section 21.1.

3.3 Randomization

On Day -1 eligible subjects who meet inclusion and do not meet exclusion criteria will be randomized in a 1:1:1 ratio to receive 1600 mg ACP-044, administered either once daily or 400 mg every 6 hours, or placebo, according to a randomization schedule. Randomization will be stratified by site.

Approximately three sites in the United States (US) will screen approximately 375 subjects and approximately 240 subjects are planned for randomization (80 subjects per treatment group).

3.4 Blinding

Treatment assignments will be blinded to all study subjects, Investigators, site personnel (with the exception of the unblinded pharmacist and verifier), and Sponsor personnel. An unblinded pharmacist and a verifier will prepare the study drug and shall have no role in pain assessment or patient care.

In order to maintain the double-blind design of the study, all subjects will receive the same numbers of tablets at the same intervals (four times per day), regardless of the treatment group to which they are randomized. Placebo tablets will be size- and color-matched to the ACP-044 tablets.

Table 1 Blinded Daily Dosing Regimen

	First Daily Dose	Second Daily Dose	Third Daily Dose	Fourth Daily Dose
Placebo	PPPP	P	P	P
ACP-044 400 mg (every 6 hours)	XPPP	X	X	X
ACP-044 1600 mg (once-daily)	XXXX	P	P	P

Abbreviations: P=ACP-044 placebo table; X=ACP-044, 400 mg tablet

3.5 Determination of Sample Size

The planned sample size is 240 randomized subjects (80 subjects per treatment group).

Assuming a treatment difference in the AUC from 0 to 24 hours is 23 points between the ACP-044 groups and placebo group, and the common standard deviation is 50 points, 76 subjects per treatment arm will provide approximately 80% power to detect the assumed treatment difference using a t-test between each ACP-044 dose arm and placebo at a 2-sided significance level of 0.05. Assuming no more than 5% of randomized subjects will be excluded from primary efficacy analysis, 80 subjects per treatment group will be randomized.

3.6 Coronavirus Disease 2019

Mechanisms to record information on the potential impact of the COVID-19 pandemic on data itself, as well as data collection and integrity, were implemented (as detailed in the "Guidance for Site Documentation and Data Management Querying of Data Impacted by COVID-19" [GSD]).

The impact of COVID-19 on the statistical analysis is discussed in Section 6 Subject Disposition, Section 7 Protocol Deviations, Section 11 Concomitant and Post-Treatment Medication, and Section 14.1 Adverse Events.

4 ANALYSIS SETS

Randomized Analysis Set

The Randomized Analysis Set includes all randomized subjects. Subjects will be analyzed based on their randomized treatment assignment.

Safety Analysis Set

The Safety Analysis Set includes all randomized subjects who received at least one dose of study drug (ACP-044 or placebo). Subjects will be analyzed based on the treatment that they actually received. The Safety Analysis Set will be used for all safety analyses.

Full Analysis Set

The Full Analysis Set includes all randomized subjects who received at least one dose of study drug post operation and who have undergone bunionectomy. Subjects will be analyzed based on their randomized treatment assignment. The Full Analysis Set will be used for the analysis of all efficacy endpoints.

Per-protocol Analysis Set

The Per-protocol (PP) Analysis Set includes all subjects in the Full Analysis Set who did not have any protocol deviations that could potentially have a substantial impact on the primary efficacy outcome. The precise reasons for excluding subjects from the PP Analysis Set will be fully defined and documented in a separate PD log file prior to the clinical database lock. Subjects will be analyzed based on their randomized treatment assignment. The PP Analysis Set will be used for supportive analyses of the primary efficacy endpoint.

Pharmacokinetics Analysis Set

The Pharmacokinetics Analysis Set includes all subjects with at least one measurable ACP-044 plasma concentration.

5 DATA HANDLING CONVENTIONS

5.1 General Data Reporting Conventions

Continuous variables will be summarized using the following descriptive statistics: number of subjects, mean, median, standard deviation (SD), standard error (SE), minimum, and maximum. Unless specified otherwise, means, medians, and confidence intervals (CIs) will be presented to one more decimal place than the raw data, and the standard deviations and standard errors will be presented to two more decimal places than the raw data. In general, the maximum number of decimal places is 4 and values will be truncated to 4 decimal places in situations where there are more than 4 decimal places. Wherever possible data will be decimal aligned.

Categorical variables will be summarized by the number of subjects and the percent of subjects in each category; the number of subjects and the percentage of subjects with missing data will be summarized for demographic and baseline characteristics (if applicable). Categories with zero counts will not have zero percentages displayed. For demographic summaries, percentages will be calculated by using the total number of subjects in the given treatment group as the denominator. Percentages will be presented with one decimal place.

Unless specified otherwise, all statistical hypothesis tests will be 2-sided and performed at the significance level of 5% for treatment comparisons and all CIs will be 2-sided 95% CIs. P-values will generally be presented to 4 decimal places; values less than 0.0001 will be presented as <0.0001.

Values that are collected with "<" or ">" signs will generally be analyzed as the numerical value without the sign in tables and figures. In listings, these data will be reported as collected with the sign.

5.2 Derived Variables

In general, the assessment scale scores will be directly taken from the electronic case report form (eCRF), the variables of mean area under the curve values will be derived and used for all analyses.

5.2.1 Numeric Rating Scale (NRS) of Pain Intensity Scores

A 0-10 NRS will be used to assess the subject's pain intensity scores from time 0 (when first dose on Day 1 is administered) through the subsequent 72 hours, and for the 7 days after EOT, subjects will be asked: "Please describe your foot pain at the present time from 0 to 10 where "0" means "no pain at all" and "10" means "the worst pain imaginable".

Pain (efficacy) assessments are to be recorded hourly (±15 minutes) for a 12-hour interval on Day 1 beginning immediately prior to the fifth dose (time 0) administered in the study.

Pain assessments will continue every 3 hours during the subsequent 12 hour interval, and then every 6 hours thereafter until 72 hours, and then once daily in the morning until the follow-up clinical visit.

There will be a total of 25 scheduled timepoints across 3 days:

- Day 1 (indexed 0 to 16): 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 21, 24 hours
- Day 2 (indexed 17 to 20): 30, 36, 42, 48 hours
- Day 3 (indexed 21 to 24): 54, 60, 66, 72 hours

Missing NRS pain intensity scores will be multiply imputed by performing the following steps:

1. Subjects with missing Time 0 assessments will be imputed by using the mean of the non-missing Time 0 assessments of the same treatment group.
2. Intermediate or interval missing records, i.e. missing pain intensity scores that are between two timepoints with non-missing pain intensity scores, will not be imputed; the AUC calculation will connect the non-missing values per the trapezoidal rule (below), effectively using linear interpolation.
3. For subjects who took rescue medication the following will be performed:
    - If the subject took any rescue medication (i.e. Acetaminophen-hydrocodone oral tablet, ibuprofen), assessments within 4 hours post taking the rescue medication will be replaced by the pre-rescue value if the pain score assessed prior to starting the rescue medication is greater than the scheduled assessment. However, if a pain score assessed within the window is the same or worse than the pre-rescue value, the scheduled assessment will be retained; this process will be repeated scheduled assessments occurring within the window.
4. Right censored missing data (i.e. missing due to discontinuation) will be imputed using multiple imputation:
    - Multiple imputation will be used to impute values at the scheduled assessment times for subjects that discontinue. For the purposes of imputation, unscheduled values and pre-rescue values will not be included, but will be added back into the dataset prior to calculating the AUC; however, scheduled values will be replaced with pre-rescue values during the censoring period prior to the imputation (and solely for the purpose of imputation). First, subjects that discontinue for reasons other than lack of efficacy (LOE) and

adverse events (AEs) will have their values imputed within treatment group with MI assuming MAR with covariates for all observed time points. For subjects who discontinue due to AEs or LOE, the MI will be based off the distribution of worst values through the time of discontinuation. This will done iteratively for each time point starting at the earliest time point with missing data; covariates will be included for observed data through that time point. Full details of the MI algorithm are given in Appendix 21.2.

In general, the area under the curve (AUC) from start time of interest, t_{start}, to end time of interest, t_k, for each subject will be calculated using the linear trapezoidal method. All pain assessments (including unscheduled) will be used in the calculation.

$$AUC_{t_{start}-t_k} = \sum_{i=start}^{k-1} \frac{1}{2} * (Pain_i + Pain_{i+1}) * (t_{i+1} - t_i)$$

Where i indexes the timepoints from 0 hour to 72 hours as 0 to 24; start is the index value of the starting time; k is the index value of the end time; t_{start} is the start time of the AUC is to be calculated; t_k is the end time of the AUC to be calculated; $Pain_i$ is the NRS pain intensity score at index i; t_i is the time at index i. **Note:** The index count will be adjusted on an individual basis to account for unscheduled assessments and pre-rescue assessments. Calculations will be performed at a 1-minute precision, but final values will be scale to hours, thus the $AUC_{0\text{-}24}$ will be scaled 0-240.

For the primary efficacy endpoint the formula would be as follows:

$$AUC_{0-24} = \sum_{i=0}^{16-1} \frac{1}{2} * (Pain_i + Pain_{i+1}) * (t_{i+1} - t_i)$$

Note: The index count will be adjusted to account for unscheduled assessments.

5.2.1.1 Time Point Capping for AUC Endpoints Only

If the relative time (from time 0) of the last scheduled timepoint for the AUC is greater than or less than the nominal terminal time for the AUC endpoint being calculated, i.e. 24 hours for AUC0-24, then the calculated AUC will be scaled by the ratio of actual time to nominal time. This will ensure that the time considered for AUC will be consistent for each subject.

For example, if time 0 is at 8:00 am and the 24 hour time point is 7:50 am the next day, then the AUC will be multiplied by 1440/1430 ([nominal minutes]/[actual minutes]).

5.2.2 Time to First Rescue Medication Use

The time to first rescue medication use will be derived relative to Time 0, Day 1. The units of hours will be used for time to event analysis.

5.2.3 Amount of Rescue Medication

The average of the total amount of ibuprofen taken (400 mg per tablet), total amount of acetaminophen (325mg per tablet) and hydrocodone taken (5 mg per tablet), or instances of any rescue medication (mg) taken during 0-24, 24-48, 48-72,0-48, and 0-72 hours by treatment group will be derived. Missing amount of rescue medication data will not be imputed. Summarize out of clinic rescue tablets/usage/add mgs.

In addition, the total amount of opioids in morphine milligram equivalents (MME) taken during the time periods will be summarized by treatment group. Both in clinic and out of clinic (Diary). In clinic, subjects should only receive hydrocodone and a 1:1 conversion of hydrocodone to MME will be utilized; should subjects take any other opioid, then the CDC conversion guide will be used for converting to MME:
https://www.cdc.gov/drugoverdose/pdf/calculating_total_daily_dose-a.pdf

5.2.4 Proportion of Subject Who Did Not Use Rescue Medication

The proportion of subjects who did not use rescue medication through 24, 48, and 72 hours will be derived as the number of subjects who did not use any kind of rescue medication through 24, 48 and 72 hours divided by the number of subjects per treatment group that were still in the trial through that time point.

5.2.5 Proportion of Subjects Who Are Pain Free (NRS ≤2)

The proportion of subjects who are pain free at each scheduled time point will be derived as the number of subject who have a NRS pain intensity score ≤2 divided by the number of subjects per treatment group with non-missing NRS values. If a subject took rescue medication within 4 hours prior to the timepoint, the subject will not be considered as pain free even if the NRS pain intensity score is ≤2.

5.2.6 Number of Hours Subjects are Pain Free (NRS ≤2)

The number of hours subjects are pain free (from 0-72 hours) will be derived as follows:

A horizontal line will be drawn at NRS of 2 on a line plot of a subject's NRS pain intensity scores. Then the time between the hours of 0 and 72, inclusive, where NRS pain intensity score is less or equal to 2 will be summed; linear interpolation will be used to calculate the amount of time for when the line plot and horizontal line intersects. If a subject took rescue medication, the next 4 hours after taking any rescue medication will not be considered pain

free even if the NRS pain intensity score is ≤2. Subjects that drop out will not have the time after they drop out considered to be ≤2 following their last reported NRS.

5.2.7 Proportion of Subjects Who Are Opioid Free

The proportion of subjects who are opioid free 0-<24, 0-<48 and 0-<72 hours and 24-<48 hours and 48-72 hours will be derived as the number of subject who did not take opioids divided by the number of subjects per treatment group. Likewise, the proportion of subjects that are opioid free starting at 0, 12, 24, 36 hours will be reported. Subjects that drop out of the study prior to a given time point will be assumed to have not subsequently taken opioids, unless concomitant medication or diary data suggest that an opioid was available to them at that time (whether or not the subject reported taking it in the period in question).

5.2.8 Global Evaluation of Study Drug (GESD)

The global evaluation of study drug is assessed just before receiving the first dose of rescue medication and at the end of 24 hours, 48 hours, and 72 hours relative to time 0. This will be considered three ways:

- With all values reported as recorded
- With values following rescue censored
- With the pre-rescue values grouped with the subsequent time point and values following rescue censored

For the GESD, subjects will be asked to provide a graded response (4=excellent, 3=very good, 2=good, 1=fair, or 0=poor) to the statement "How would you rate the study drug you received to delay or reduce your foot pain?"

5.2.9 Proportion of Subjects with Severe Pain

The proportion of subjects with severe pain will be derived as the number of subjects with an NRS pain intensity score of ≥7 at any timepoint through 72 hours divided by the number of subjects per treatment group.

5.2.10 Foot and Ankle Ability Measure (FAAM)

The foot and ankle ability measure (FAAM) is assessed at Baseline and in the Safety Follow-up Period, 7 days after EOT/ET.

There are two subscales: activities of daily living (ADL) and Sports. For ADL subscale, there are 21 movements the subject rates as "4=no difficulty at all", "3=slight difficulty", "2=moderate difficulty", "1=extreme difficulty", "0=unable to do", and "N/A", and an overall grade for the level of function during usual activities of daily living, graded on a scale

of 0 to 100 with 100 being at the level of function prior to the foot and ankle problem and 0 being the inability to perform any of their usual daily activities.

For Sports subscale, there are 8 movements and an overall grade for the level of function during sports related activities, graded in the same manner as ADL (0-4 for items, 0-100 for overall grade).

Missing FAAM data will not be imputed.

Within subscale, the item score total will be derived as the sum of items with a response (NA responses are not counted). The total number of items with a response is multiplied by 4 to get the highest potential score (For ADL, if all 21 items have a response that has a score of 0-4, the highest potential score is 84, if there is one item with a NA or missing response, then the highest potential score is 80; the highest potential score for Sports subscale is derived similarly; however, the Sports subscale only has 8 items). The item score total will be divided by the highest potential score; this value will be multiplied by 100 to get a percentage. A higher percentage represents a higher level of physical function.

If there are less than 90% completed items (NA responses are considered as not completed), i.e., less than 19 items completed for ADL or less than 7 items for Sports subscales, then the corresponding item percentage score will be considered missing.

Total score will be derived as the sum of the ADL and Sports subscales. If either subscale is missing, total score will also be missing.

5.2.11 Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is assessed at Screening, Day -1 and Day 4 (EOT/ET).

The C-SSRS baseline/screening version will be administered at screening.

The C-SSRS version assessing information since the last visit will be completed at Day -1 and Day 4 (EOT/ET) visits.

There are 5 questions about suicidal ideation, representing 5 types of suicidal ideation: wish to be dead; non-specific active suicidal thoughts; active suicidal ideation with any methods (not plan) without intent to act; active suicidal ideation with some intent to act, without specific plan; active suicidal ideation with specific plan and intent. If a subject answers “yes” to any of these 5 questions, this subject will be counted as having suicidal ideation.

There are 5 questions about suicidal behavior, representing 5 types of suicidal behavior: preparatory acts or behavior; aborted attempt; interrupted attempt; actual attempt; suicide. If a subject answers “yes” to any of these 5 questions, this subject will be counted as having suicidal behavior.

Suicidality is defined as a subject who reported at least 1 occurrence of suicidal ideation or at least 1 occurrence of suicidal behavior at any post-Baseline visit including unscheduled and out of window visits.

Missing C-SSRS item scores will not be imputed.

5.3 Study Analysis Day

If the date of assessment occurs on or after the first dose date, then the study analysis day will be calculated as (date of assessment – first dose date) + 1. If the date of assessment occurs prior to the first dose date, then study analysis day will be calculated as (date of assessment – first dose post operation). There is no study analysis day 0.

5.4 Baseline Definition

In general, baseline for safety data and FAAM are defined as data collected which are prior to the administration of the first dose. If there is more than one value on or prior to first dose, the value closest to and prior to the receipt of the first dose, whether scheduled or unscheduled, will be used as the Baseline value.

NRS pain severity score does not have a baseline.

5.5 Analysis Visit Windows

Efficacy, safety, and PK assessments will be summarized by analysis visit and timepoint as presented in Table 2 and Table 3 below.

Study analysis visit window will be used to map visits using study day intervals. In Table 2, the first day of double-blind dosing is defined as “Study Analysis Day 1”.

Table 2 Analysis Visit Windows

Analysis Visit Name	Target Study Analysis Day	Study Analysis Day Interval
Baseline[1] (Day -1)	1	≤ 1 and pre-dose
Day 1	2	2
Day 2	3	3
Day 3	4	4
Day 4	5	5

[1] The Baseline is not applicable to NRS assessment.

Table 3 Analysis Timepoint Windows

Parameter	Target Study Timepoints	Timepoint Window
Vital Signs	Day -1 predose, Day 1: 2h, 4h, 6h postdose and then every 6 hours thereafter up to Day 4/EOT(12h, 18h, 24h, 30h, 36h, 42h, 48h, 54h, 60h, 66h, 72h…)	Predose assessment must be prior to dose on Day -1. For postdose records, the scheduled time will be used for analysis; the midpoint between each timepoint will be used to map the assessments such that there is no gap in windowing.
ECG	Day -1 predose, Days 1-3 at 1h postdose, and EOT/ET.	Predose assessments must be prior to dose on Day -1. For post-dose assessments, the midpoint between each timepoint will be used to map the assessments such that there is no gap in windowing.
PK Blood Samples	Day -1 predose within one hour of dosing, Day 1 between 2-3h postdose, Day 1 between 4-6h postdose, Day 2 predose, Day 3 between 2-6h postdose, Day 3 between 8-12h postdose, and in the event of an SAE or AE leading to discontinuation (if possible).	For time points prior to Day 2 predose, the window will be ±15 min on top of the collection window range. For timepoints on Day 2 and beyond will be ±1 hour on top of the collection window range.
Block Quality Assessments	Postoperatively, assessments will be performed and recorded hourly (±15 minutes) at the top of each hour until 2200 hours. Between the hours of 2200 and 0600 hours, assessments will continue to be performed hourly (±15 minutes) if subject is awake, but, even if asleep, the subject will be awakened to perform a Block Quality Assessment at 0300 hours (±15 minutes). Block Quality Assessments will continue until the block is removed.	For each hour 0-24 post surgery completion, the time of the surgery completion defines analysis timepoint 0, a +/- 30 min window will be used. For time 0 the window will start at the time of surgery completion.
NRS	Day 1: 0h (Time 0), 1h, 2h, 3h, 4h, 5h, 6h, 7h, 8h, 9h, 10h, 11h, 12, 15h, 18h, 21h, 24h, 30h, 36h, 42h, 48h, 54h, 60h, 66h, 72h, and then once daily in starting the morning of Day 4 until the follow-up clinic visit.	For each timepoint up to 72h the window will be ±30 min. For AUC calculation, all NRS timepoints will be considered. Only windowed assessments will be summarized in the by timepoint summary and for specific steps outlined for the multiple imputation of missing data in the sensitivity analysis.

5.5.1 Unscheduled Assessments

Both scheduled and unscheduled assessments, including the assessments at early termination visits, will be considered for planned timepoint summaries based on the above analysis visit windowing rules. All assessments will be presented in data listings.

5.5.2 Multiple Measurements within Visit or Timepoint Windows

In general, in the event that more than one assessment falls within a given visit or timepoint window, the assessment closest to the target study day and time will be selected for the

by-visit or by-timepoint analysis. If two assessments are equidistant from the target study day or nominal time, then the chronologically last assessment will be used. Details are provided in a separate programming conventions document.

All NRS timepoints will be used regardless if the assessment is scheduled or unscheduled.

For safety analyses where the most extreme values should be selected, e.g. overall post-baseline minimum, overall post-baseline maximum, and potentially clinically important values for overall post-Baseline summaries, all non-missing post-Baseline values should be considered, regardless of whether the value is selected for the by-visit summaries. All results will be presented in data listings.

5.6 Missing or Incomplete Date for Last Dose of Study Drug

For subjects with completely missing last dose date, the last dose date will be imputed by the last expected dosing date, defined as the earliest of the following dates: last dose administer date + scheduled dosing interval per protocol, EOT/ET date. For subjects with partial missing last dose date, the imputation will be compared against the last expected dosing date as defined above. Detailed algorithms will be documented in a separate programming specifications document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.7 Missing or Incomplete Dates for Concomitant or Post-Treatment Medications

Missing or incomplete medication start or end dates will be imputed for the purpose of determining whether the medications are taken concomitantly (see Section 11 for definition). When the chronological order of medication use relative to the study drug treatment period is unclear due to missing or incomplete date(s), the medication will be considered as concomitant. The imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.8 Missing or Incomplete Dates for Adverse Events

Missing or incomplete adverse event (AE) start dates will be imputed for the purpose of determining whether the AEs are treatment-emergent (see Section 14.1 for definition). When the chronological order of an AE onset relative to the study drug treatment period is unclear due to missing or incomplete date(s), the AE will be considered as treatment-emergent. The imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographics and Baseline characteristics will be summarized for the Safety Analysis Set using descriptive statistics. Variables include age, sex, primary race, ethnicity, height, weight, and body mass index (BMI).

Race will also be categorized by an additional binary category of White vs. Non-White. The eCRF reported age will reflect a subject's age at the Screening visit date; age at screening will be used for all analysis.

6.1 Block Quality Assessments

Assessments of the popliteal block depth (Block Quality Assessment) will be summarized by timepoint (For each hour 0-24 post surgery completion) and treatment group. The assessment consists of three categories, motor, sensory and pain. The categories of motor and sensory will be categorically summarized as None, Some, Normal. The pain category will consist of the intervals based off of a categorical NRS pain intensity score: (0= NRS 0-2), (1=NRS 3-6) and (2=NRS 7-10). Each component of the block quality assessment collected will be listed.

6.2 Missing Severity Assessment for Adverse Events

If will be used for incidence summaries, and the actual values will be presented in data listings.

6.3 Missing Relationship to Study Drug for Adverse Events

If the relationship to study drug is missing for a treatment-emergent AE, a causality of "Related" will be assigned. The imputed values for relationship to study drug will be used for incidence summaries, and the actual values will be presented in data listings.

6.4 Character Values of Clinical Laboratory Variables

If the reported value of a clinical laboratory variable cannot be used in a summary due to, for example, a character string reported for a numeric variable, an appropriately determined coded value will be used in the summary. The coding algorithms will be detailed in the analysis dataset specification document. The actual values as reported in the database will be presented in data listings.

7 SUBJECT DISPOSITION

For subjects who participate in the screening phase but are not randomized (screen failures), their demographics information (including age, sex, and primary race), screen failure reasons (the specific inclusion/exclusion criterion (or criteria) not met or other reasons including the reason due to the COVID-19 public health emergency (PHE) and protocol version will be listed. If a subject is re-screened, then the re-screening subject ID and the final enrollment

status (whether eventually enrolled) will also be displayed in this listing. In addition, the frequency that the screen failure reasons are reported will be summarized. Note that one subject may be deemed ineligible for multiple inclusion/exclusion criteria and may be allowed to rescreen with the permission of the Medical Monitor, provided the screen failure was due to a temporary condition that subsequently resolved.

The number of sites that screened at least 1 subject, number of sites that randomized at least 1 subject, number of subjects screened, and number of unique subjects screened will be tabulated. In addition, the number of subjects enrolled at each site will also be tabulated by Analysis Set and by treatment group and overall.

The number and percentage of subjects who completed the study, discontinued early (all discontinued and by discontinuation reasons including reason due to the COVID-19 PHE), and the reason for discontinuation will be summarized by treatment group using the Randomized Analysis Set. All subjects excluded from the Safety, Full or Per-protocol Analysis Sets, and the reason(s) for exclusion will be listed. The number and percentage of subjects who are excluded from the Per-protocol Analysis Set will be presented in a summary table by reason, and by treatment group and overall.

8 PROTOCOL DEVIATIONS

Protocol deviations will be reviewed periodically over the course of the study. The review process, definition of the deviation categories, and the classification of a deviation as major or minor are detailed in the Protocol Deviation Management Plan. Protocol deviations will also be assessed with respect to relationship to the COVID-19 PHE. A summary of the number and percentage of subjects with major protocol deviations for each deviation category may be presented by treatment group for the Randomized Analysis Set in three ways: all major protocol deviations, COVID-19-PHE related major protocol deviations (if applicable), and non-COVID-19-PHE related major protocol deviations (if applicable). Three data listings of all protocol deviations, COVID-19-PHE related protocol deviations (if applicable), and non-COVID-19-PHE related protocol deviations (if applicable) may be provided.

9 MEDICAL HISTORY

Medical history data will be coded using Medical Dictionary for Regulatory activities (MedDRA) version 24.0. The subject incidence will be summarized for each system organ class (SOC) and preferred term for the Safety Analysis Set. A subject will be counted only once per SOC or per preferred term for the summary.

A listing of the SOC, preferred term, verbatim term for the medical history condition/event, start and stop dates (when available), and an indicator for whether or not the condition is ongoing will be provided.

10 EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

Summaries of exposure and compliance to study drug will be provided for the Safety Analysis Set.

10.1 Exposure to Study Drug

Exposure in days, the number of doses administered, the number of missed doses and the number of total tablets dosed will be summarized.

10.2 Treatment Compliance

Study drug compliance will be calculated as the number of tablets taken, divided by the expected number of tablets taken*100.

Treatment compliance will be summarized as a categorical variable. The number and percentage of subjects within each of the following compliance levels will be tabulated: <80%, 80 to 120% and >120%.

The number of missed doses will be summarized as a categorical variable: none, at least 1, 2-4, 5-8, 8-12, and >13.

11 CONCOMITANT AND POST-TREATMENT MEDICATION

Prior medication is defined as any medication with stop dates prior to the date of the first dose of study drug. Concomitant medication is defined as any medications that are ongoing at the first dose of study drug or with a start date between the dates of the first and last doses of study drug, inclusive. Post-treatment medication is defined as any medication with a start date after the date of the last dose of study drug. Medications will be coded using WHO Drug Global Dictionary March 2021 or newer version.

The number and percentage of subjects taking prior, concomitant and post-treatment medications will be tabulated separately by each drug class (ATC Level 3) and medication preferred term, treatment group and overall for Safety Analysis Set. Multiple medication usage by a subject in the same category will be counted only once.

Summary of pharmacokinetic concomitant medications will be summarized. Pharmacokinetic concomitant medication is defined as any medications that are ongoing at first dose of study drug or with a start date between the dates of the first and last doses of study drug, inclusive; and taken within two days prior to a pharmacokinetic sample collection. The details of this summary will be discussed in a separate PK analysis plan.

Listings of the prior, concomitant and post-treatment medications will also be provided.

COVID-19 Public Health Emergency (PHE) Related Medications

Concomitant medication analyses described above may also be tabulated and listed by relationship to the COVID-19 PHE (Not related to the COVID-19-PHE vs. Related to the COVID-19-PHE), if applicable.

12 EFFICACY ANALYSES

Unless otherwise specified, all efficacy analyses will be performed using the planned treatment assignments based on the randomization schedule for the Full Analysis Set.

12.1 Efficacy Variables

Primary Efficacy Endpoint

- The mean AUC_{0-24} of NRS of pain intensity scores (using 4 hour window for rescue medications)

Key Secondary Efficacy Endpoints

- Time to first rescue medication use after time 0 (when the first dose on Day 1 is administered)
- Proportion of subjects who were opioid free through 24 hours
- Proportion of subjects who were opioid free through 48 hours
- Proportion of subjects who were opioid free through 72 hours

Secondary Efficacy Endpoints

- The mean AUC_{0-48} of NRS of pain intensity scores (using 4 hour window for rescue medications)
- The mean AUC_{0-72} of NRS of pain intensity scores (using 4 hour window for rescue medications)
- The mean AUC_{0-4} of NRS of pain intensity scores (using 4 hour window for rescue medications)
- The mean AUC_{0-6} of NRS of pain intensity scores (using 4 hour window for rescue medications)
- The mean AUC_{0-12} of NRS of pain intensity scores (using 4 hour window for rescue medications)
- The mean AUC_{24-48} of NRS of pain intensity scores (using 4 hour window for rescue medications)

- The mean AUC_{48-72} of NRS of pain intensity scores (using 4 hour window for rescue medications)
- Amount of rescue medication taken during 0-24, 24-48, and 48-72 hours (individually), 0-48 hours, and 0-72 hours
- Proportion of subjects who did not use rescue medication through 24, 48, and 72 hours
- Proportion of subjects who are pain free (NRS ≤2) at 24, 48, and 72 hours
- Number of hours subjects are pain free (NRS ≤2)
- Proportion of subjects who were opioid free during 24-48 hours and 48-72 hours
- Global evaluation of study drug just before time of first rescue and at the end of 24, 48, and 72 hours relative to the time 0 (when the first dose on Day 1 is administered)

Exploratory Efficacy Endpoints

- Proportion of subjects with severe pain (defined as an NRS pain intensity score ≥7 at any timepoint through 72 hours)
- Foot and Ankle Ability Measure (FAAM)

12.2 Adjustment for Covariates

Unless otherwise specified, there will be no adjustment for covariates.

12.3 Multiple Comparisons / Multiplicity

A fixed sequence method at alpha=0.05 for the primary and key secondary efficacy endpoints will be used to control for multiplicity. The following will be the sequence for testing:

1. 400 mg Q6H compared to placebo on mean AUC_{0-24} of NRS of pain intensity scores
2. 1600 mg QD compared to placebo on mean AUC_{0-24} of NRS of pain intensity scores
3. 400 mg Q6H compared to placebo on time to first rescue medication use after time 0
4. 400 mg Q6H compared to placebo on proportion of subjects who were opioid free through 24 hours
5. 1600 mg QD compared to placebo on time to first rescue medication use after time 0
6. 1600 mg QD compared to placebo on proportion of subjects who were opioid free through 24 hours
7. 400 mg Q6H compared to placebo on proportion of subjects who were opioid free through 48 hours

8. 1600 mg QD compared to placebo on proportion of subjects who were opioid free through 48 hours
9. 400 mg Q6H compared to placebo on proportion of subjects who were opioid free through 72 hours
10. 1600 mg QD compared to placebo on proportion of subjects who were opioid free through 72 hours

The testing procedure will stop once a p-value in the sequence is greater than 0.05. P-values that are less than or equal to 0.05 prior to stopping, will be declared statistically significant; p-values after stopping will be nominal. P-values obtained from the stratified log-rank test stratified by site will be used to for the endpoints that compare the time to first rescue medication use after time 0 by treatment group.

12.4 Examination of Subgroups

Treatment effect will be examined with respect to the primary endpoint using ANOVA described in Section 13.1.1 for the following subgroups:

- Age (<median or ≥median)
- Sex (male or female; reporting on the male subgroup may be omitted if there are insufficient counts)
- Primary race (white or non-white)

13 METHODS OF EFFICACY ANALYSES

In general, comparisons will be pairwise between ACP-044 400 mg Q6H, ACP-044 1600 mg QD and placebo.

13.1 Primary Efficacy Endpoints

13.1.1 Primary Analysis

The primary clinical question of interest for the primary objective is: what is the difference in the mean $AUC_{0\text{-}24}$ of the NRS of pain intensity scores comparing ACP-044 400 mg Q6H vs. placebo and ACP-044 1600 mg QD vs. placebo, in subjects with acute postoperative pain following orthopedic surgery?

The estimand is described by the following attributes:

Population: Adult subjects that have acute postoperative pain following orthopedic surgery, as defined by the inclusion/exclusion criteria of the study.

Variables (Primary Endpoint): mean $AUC_{0\text{-}24}$ of NRS pain intensity scores

Treatment Condition: ACP-044 400 mg Q6H, ACP-044 1600 mg QD and placebo, with use of rescue medications.

Intercurrent Events and Strategies:

- Treatment Discontinuation will be addressed by the hypothetical strategy, i.e., assuming that subjects with these intercurrent events (ICEs) continue to experience the following the distribution of worst observed pain prior to the ICE occurrence if they discontinue for lack of efficacy or an adverse event; subjects discontinuing for other reasons will be assumed to follow the distribution of values within their treatment group.
- Taking Rescue Medication will be addressed by the composite strategy of estimating pain scores in the presence of taking rescue medication by replacing assessments taken within a 4-hour window with the pre-rescue value. Should the NRS pain assessment taken immediately prior to rescue be less than a scheduled pain assessment, that scheduled pain assessment will be retained.

Alternative approaches to handling ICEs will be addressed in sensitivity analyses.

Population-level Summary Measure:

The treatment difference (ACP-044 400 mg Q6H vs placebo, ACP-044 1600 mg QD vs placebo) in the mean NRS pain intensity score $AUC_{0\text{-}24}$.

Hypotheses

The primary endpoint is $AUC_{0\text{-}24}$ based on the pain intensity scores from 0 to 24 hours assessed using NRS.

Let Δ_{auc_1} and Δ_{auc_2} be the difference between each of the two ACP-044 dose groups and placebo group in the mean $AUC_{0\text{-}24}$, respectively.

The null hypotheses are: $\Delta_{auc\_i} = 0$ and the alternative hypothesis is: $\Delta_{auc\_i} \neq 0$, i=1,2.

Primary Estimator

The hypothesis testing will be performed for the Full Analysis Set using analysis of variance (ANOVA) with treatment group and site as factors in the model. Prior to analysis, missing pain assessments will be imputed before calculating AUC per Section 5.2.1. In addition, the AUC will have the last time point adjusted by as described in Section 5.2.1.1.

The ANOVA results will be processed using the SAS MIANALYZE procedure to yield a combined estimate for treatment effect, standard error, and associated 95% confidence interval (CI) and p-value (Rubin 1976).

The treatment effect size, Cohen's d, will be presented for both ACP-044 comparisons to placebo using the following formula:

$$Effect\ Size = \frac{LS\ mean\ difference}{Standard\ Deviation}$$

Where Standard Deviation is approximated:

$$Standard\ Deviation = \frac{Standard\ Error\ of\ LS\ mean\ Difference}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

And n_1 and n_2 are the sample sizes in each group. The sign (+ or -) of the effect size will be chosen so that a positive value favors ACP-044.

Summary statistics for the mean NRS pain intensity assessments (before and after single imputation) and the mean of AUC of pain intensity scores, the LS means, the between-group difference in LS mean with the corresponding confidence interval, p-value, and the effect size (Cohen's d) will be presented.

All pairwise treatment group comparisons will be presented.

A line plot of the mean (+/- the SE) NRS pain intensity assessments over time will be presented by treatment group.

13.1.2 Sensitivity Analyses

The following sensitivity analyses of the primary efficacy endpoint are planned to describe different strategies for intercurrent events.

13.1.2.1 Alternative Handling of Rescue Medication Censoring

As an alternative strategy to the intercurrent event of taking rescue medication, the impact of the censoring described in Section 5.2.1 will be investigated a few different ways:

1. A 2 hour censoring window will be used.
2. A 6 hour censoring window will be used.
3. No values will be censored for use of rescue medication.

For each of the above scenarios, the other components of the imputation and data handling will be identical to the primary and the NRS pain intensity score $AUC_{0\text{-}24}$ values will then be calculated, and analyzed similarly as in the primary analysis.

13.1.2.2 Multiple Imputation Variations

As a sensitivity to the primary endpoint analysis, the multiple imputation approach will be adjusted to better understand the impacts of its assumptions. Four approaches will be investigated:

1. All subjects, regardless of reason for withdrawal, will have missing data imputed under a missing-at-random assumption within treatment groups.
2. All subjects, regardless of reason for withdrawal, will have missing data imputed using the distribution of the worst values through that point.
3. Subjects that withdraw due to "other" and "unknown" reasons will be grouped with those that withdraw for LOE and AEs; all other imputation rules from the primary analysis will be unchanged.
4. In addition, the placebo-based imputation will be applied to the missing data due to AE or LOE. Like "standard" multiple imputations, except parameters for imputation model will be obtained from only the placebo (control) group. Both placebo and drug groups will be imputed based on the imputation model derived from placebo data. If drug improves outcomes prior to dropout, this benefit is carried into subsequent imputed values, but will diminish over time in accordance with the correlation structure. Subjects that discontinue for reasons other than LOE or AE will have their values imputed within treatment group with MI assuming MAR with covariates for all observed time points.

13.1.3 Supportive Analyses

Supportive analysis of the primary efficacy endpoint **may** be analyzed in a similar fashion as the primary analysis, using the Per-Protocol analysis set.

13.2 Key Secondary Efficacy Endpoints

Analysis of the key secondary efficacy endpoints will be performed using the Full Analysis Set.

Time to first rescue medication use after time 0 (hours)

A stratified log-rank test stratified by site will be used to compare the time to first rescue medication use after time 0 by treatment group. The follow-up time will end at 72 hours. The time at which subjects discontinue due to the lack of efficacy will be considered as an event. Subjects who do not take rescue medication and who do not discontinue for lack of efficacy will be censored at 72 hours or their last follow-up time, whichever is earlier. The estimate and 95% CI's for the median time to event will be summarized by treatment group, where

event is the time to first rescue medication use after time 0. The p-value of the stratified test of equality over treatment group for each pairwise comparison will be summarized.

Sample SAS code:

```
proc lifetest data=indata;
time time*cnsr(1);
strata site /group=trt;
run;
```

where time is the time to first rescue or censor, cnsr is the censor status, site is the site variable, and trt is the treatment group variable.

In addition, a stratified by site Cox proportional hazards model treatment group as a factor will be used to analyze the time to rescue medication or censoring. The follow-up time will end at 72 hours. The time at which subjects discontinue due to the lack of efficacy will be considered as an event. Subjects who do not take rescue medication and who do not discontinue for lack of efficacy will be censored at 72 hours or their last follow-up time, whichever is earlier.

The hazard ratio and 95% confidence interval will be summarized for each pairwise comparison as defined previously.

Kaplan-Meier curves of time to first rescue will be presented.

Sample SAS code:

```
proc phreg data=indata covsandwich;
class trt(ref="Placebo");
model time*cnsr(1)= trt/ ties=exact;
strata site;
run;
```

In addition, the number and percentage of subjects taking rescue medication or discontinuing due to lack of efficacy and the number and percentage of subjects censored will be summarized.

Proportion of subjects who were opioid free through 24, 48, and 72 hours

The proportion of subjects who are opioid free (did not take acetaminophen-hydrocodone oral tablet as rescue medication or any other opioid) through 24, 48, and 72 hours and during 0-<24 hours, 24-<48 hours, and 48-72 hours will be analyzed using a stratified CMH test controlling for site.

The risk difference, 95% confidence interval and p-values will be provided by treatment group pairwise.

13.3 Secondary Efficacy Endpoints

Analyses of the secondary efficacy endpoints will be performed using the Full Analysis Set.

The following secondary efficacy endpoints will be analyzed similarly as the primary analysis for the primary efficacy endpoint.

- The mean $AUC_{0\text{-}48}$ of NRS of pain intensity scores

- The mean AUC_{0-72} of NRS of pain intensity scores
- The mean AUC_{0-4} of NRS of pain intensity scores
- The mean AUC_{0-6} of NRS of pain intensity scores
- The mean AUC_{0-12} of NRS of pain intensity scores
- The mean AUC_{24-48} of NRS of pain intensity scores
- The mean AUC_{48-72} of NRS of pain intensity scores

Amount of rescue medication taken during 0-24, 24-48, 48-72, 0-48, and 0-72 hours

The following will be summarized:

- The number of tablets of ibuprofen taken during 0-24, 24-48, 48-72, 0-48, and 0-72 hours.
- The number of tablets acetaminophen-hydrocodone taken during 0-24, 24-48, 48-72, 0-48, and 0-72 hours.
- The number of instances of any rescue medication taken during 0-24, 24-48, 48-72, 0-48, and 0-72 hours.
- The amount of opioids in Morphine Milligram Equivalents (MME) taken during 0-24, 24-48, 48-72, 0-48, and 0-72 hours

Each of the above items will be analyzed using ANOVA with treatment group and site as factors in the model by rescue medication type.

The LS mean (SE), 95% confidence intervals, and p-values will be presented by treatment group. The LS mean difference (SE), associated 95% confidence intervals, p-values, and Cohen's d will be presented by treatment group pairwise.

Proportion of subjects who did not use rescue medication through 24, 48, and 72 hours

Given the in-clinic setting of this study, subjects with no records of rescue medication dispensation will be assumed to not have used any rescue medication.

The proportion of subjects who did not use rescue medication (neither acetaminophen-hydrocodone oral tablets nor ibuprofen nor any other) through 24, 48, and 72 hours will be summarized and will be analyzed using a stratified Cochran-Mantel-Haenszel (CMH) test controlling for site.

The risk difference, 95% confidence interval and p-values will be provided by treatment group pairwise.

Proportion of subjects who were opioid free during 24-48 hours and 48-72 hours

The proportion of subjects who are opioid free (did not take acetaminophen-hydrocodone oral tablet as rescue medication or any other opioid) through 24, 48, and 72 hours and during 0-<24 hours, 24-<48 hours, and 48-72 hours will be analyzed using a stratified CMH test controlling for site.

The risk difference, 95% confidence interval and p-values will be provided by treatment group pairwise.

Proportion of subjects who are pain free (NRS ≤2) at 24, 48, and 72 hours

The proportion of subjects who are pain free at 24, 48, and 72 hours will be analyzed using a stratified CMH test controlling for site.

The risk difference, 95% confidence interval and p-values will be provided by treatment group pairwise.

Number of hours subjects are pain free (NRS ≤2)

The number of hours subjects are pain free will be analyzed using ANOVA with treatment group and site as factors in the model. Each active dose group will be compared to placebo using this model.

The LS mean (SE), 95% confidence intervals, and p-values will be presented by treatment group. The LS mean difference (SE), associated 95% confidence intervals, p-values, and Cohen's d will be presented by treatment group pairwise.

Global evaluation of study drug just before time of first rescue and at the end of 24, 48, and 72 hours

The global evaluation of study drug (GESD) will be summarized as a categorical and as a continuous variable for each visit and for each censoring approach:

- With all values reported as recorded
- With values following rescue censored
- With the pre-rescue values grouped with the subsequent time point and values following rescue censored

The GESD will be analyzed by visit using ANOVA with treatment group and site as factors in the model.

The LS mean (SE), 95% confidence intervals, and p-values will be presented by treatment group. The LS mean difference (SE), associated 95% confidence intervals, p-values, and Cohen's d will be presented by treatment group pairwise.

13.4 Exploratory Efficacy Endpoints

Proportion of subjects with severe pain (defined as an NRS pain intensity score ≥7 at any timepoint through 72 hours)

The proportion of subjects with severe pain will be analyzed using a stratified CMH test controlling for site.

The odds ratio, 95% confidence interval and p-values from the CMH test will be provided by treatment group pairwise.

Foot and Ankle Ability Measure (FAAM)

The summary of each movement category will be summarized descriptively as a categorical variable. The overall grades (ADL item score percentage, patient reported ADL overall score, Sports item score percentage, patient reported Sports overall score, and derived total score) will be summarized as a continuous variable. The change from baseline in overall grade will be analyzed using analysis of covariance (ANCOVA) with treatment group and site as factors and baseline (Day -1) FAAM as a covariate in the model.

The LS mean (SE), 95% confidence intervals, and p-values will be presented by treatment group. The LS mean difference (SE), associated 95% confidence intervals, p-values, and Cohen's d will be presented by treatment group pairwise.

14 SAFETY ANALYSES

All safety analyses will be performed using the actual treatment received for the Safety Analysis Set.

14.1 Adverse Events

Adverse events will be coded using MedDRA dictionary, Version 24.0 or newer.

An AE (classified by preferred term) will be considered a treatment-emergent AE (TEAE) if started after first dose administration and no later than last dose date + 30 days. AEs reported on Day -1 based on Baseline (pre-dose) findings (e.g., clinically significantly abnormal vital signs, laboratory test results, or electrocardiogram parameters) will not be considered as TEAEs.

The event counts, the number, and percentage of subjects reporting TEAEs in each treatment group will be tabulated by system organ class (SOC) and preferred term; by SOC, preferred term, and maximum severity; and by SOC, preferred term, and relationship to study drug. If more than one TEAE occurs with the same preferred term for the same subject, then the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to study drug. The display in these tables will be sorted alphabetically by SOC and then by descending subject frequency for the preferred terms (counts from ACP-044 treated group) within each SOC.

The event counts, the number and percentage of subjects with any TEAEs will also be tabulated by preferred term without SOCs. This table will be sorted by descending subject frequency within ACP-044 treated group.

The incidence of most frequently reported (preferred terms reported by ≥ 5% of subjects in any treatment group) TEAEs, SAEs reported after treatment start, TEAEs leading to discontinuation of study drug, and TEAEs related to study drug will be summarized by SOC, preferred term, and treatment group. The tables will be sorted alphabetically by SOC and then by descending frequency within each SOC in ACP-044 treated group. In addition, the incidence of fatal treatment-emergent AEs (i.e., events that cause death) will be summarized separately by SOC and preferred term.

The incidence of adverse events of special interest, such as opioid-related adverse events (ORAEs) will be summarized separately by SOC and preferred term. ORAEs evaluation will be based on the following prespecified preferred terms of nausea, vomiting, constipation, pruritus, pruritus generalized, somnolence, sedation, respiratory depression, urinary retention hypotension, pre-syncope, syncope, and bradycardia, regardless of whether a subject actually received an opioid medication.

These summary tables except for the most frequently reported TEAEs tables may also be presented for COVID-19 and non-COVID-19 related events, if applicable.

An AE listing by subject will display all events, including those which occur during screening, and will include the verbatim term in addition to the MedDRA SOC and preferred term. This listing will also include all relevant eCRF data associated with the event: date of onset, date resolved, date of last dose, severity, frequency, outcome, relationship to study drug, and action taken with study drug. A listing by subject will display all TEAE, excluding AE that are not treatment-emergent. Separate listings will be presented for subjects with treatment-emergent SAEs, related TEAEs, TEAEs leading to discontinuation, fatal TEAEs (if any), and COVID-19 related TEAEs (if applicable).

14.2 Clinical Laboratory Variables

Clinical laboratory tests are performed at Day -1 predose and Day 4 (EOT/ET).

The laboratory evaluations will include, but are not limited to, the following:

- Clinical chemistry serum tests (CHEM)
    - Sodium (Na), potassium (K), chloride (Cl), phosphorus (P), calcium (Ca),carbon dioxide (CO2), blood urea nitrogen (BUN), creatinine (CR), uric acid,
        - Mg should only be performed at Visit 1 (Screening)
    - Estimated glomerular filtration rate (eGFR) should only be performed at Visit 1 (Screening)

  - Alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), total bilirubin (TBIL), lactate dehydrogenase (LDH)
  - Creatine kinase (CK)/creatine phosphokinase (CPK)
  - Lipid panel should only be performed at Visit 1 (Screening):
    - Total cholesterol, HDL-cholesterol, triglycerides, LDL-cholesterol, cholesterol/HDL ratio, non-HDL cholesterol; should only be performed at Visit 1 (Screening)
  - HbA_{1c} should only be performed at Visit 1 (Screening)
  - Glucose
  - Albumin (ALB) should only be performed at Visit 1 (Screening)
  - Total protein should only be performed at Visit 1 (Screening)
- Hematology tests
  - Complete blood count (CBC) including:
    - White blood cell (WBC) count
    - Complete differential (relative and absolute)
    - Hematocrit (Hct), hemoglobin, red blood cells (RBC), platelets
    - Reticulocyte count
- Endocrinology
  - Thyroid stimulating hormone (TSH) and free T4 (Screening)
  - Vitamin B12 should only be performed at Visit 1 (Screening)
- Pregnancy test
  - A serum pregnancy test should only be performed at screening for women of childbearing potential
  - A urine pregnancy test should be performed before surgery on Day -1 for women of child-bearing potential
- COVID-19 test:
  - COVID-19 diagnostic PCR test at Screening
  - COVID-19 rapid antigen test on Day-1 predose

- Urinalysis (UA)
    - Blood, RBCs, WBCs, protein, glucose, ketones, specific gravity, pH, nitrates
- Urine toxicology screen
    - Urine toxicology screen will test for controlled substances. The following controlled substances may be tested with a urine toxicology screen at Screening and Day -1: amphetamine, barbiturates, benzodiazepines, cocaine, methadone, morphine/opiates, methamphetamine, marijuana (THC), phencyclidine (PCP), ecstasy (MDMA).

Clinical laboratory values (in Système International [SI] units) and the change from Baseline values will be summarized by treatment group at each post-Baseline visit using descriptive statistics. The overall minimum, maximum as well as the last post-Baseline observed and change from Baseline values will also be summarized. For urinalysis with categorical results, the number and percentage of subjects will be tabulated by category at Baseline and each post-Baseline visit, and the denominator is the number of subjects with non-missing values for the given parameter, visit and treatment group.

Laboratory values will also be summarized in shift tables by treatment group, to determine the number and percentage of subjects with values classified as below, within, and above normal ranges at each post-Baseline visit relative to the same classification at the Baseline visit. For the by-visit shift summary, the denominator is the number of subjects with non-missing values at Baseline and the given visit for the given parameter and treatment group. For the shift to the overall post-Baseline minimum or maximum, all post-Baseline values will be considered, including unscheduled and out of window values and the denominator is the number of subjects with non-missing Baseline value and at least 1 post-Baseline value for the given parameter and treatment group.

Clinical laboratory values are potentially clinically important (PCI) if they meet either the low or high PCI criteria listed in Tables 4 and 5. The number and percentage of subjects with post-Baseline PCI values for each of the categories in Table 4 and 5 will be summarized by treatment group for selected parameters. For the overall post-Baseline summaries of PCI values, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the overall post-Baseline summary, the numerator of the percentage is the number of subjects with at least 1 post-Baseline PCI laboratory value for the given parameter and treatment group, and the denominator is the number of subjects with at least 1 post-Baseline laboratory value for the given parameter and treatment group.

Table 4 Criteria for Potentially Clinically Important Laboratory Values – Hematology and Chemistry

Analyte	Conventional Unit	Low PCI Criteria	High PCI Criteria	SI Unit	Low PCI Criteria	High PCI Criteria
Hematology (whole blood)						
Hemoglobin (male)	g/dL	<11	>18	g/L	<110	>180
Hemoglobin (female)	g/dL	<10	>17	g/L	<100	>170
Hematocrit (male)	%	<30	>55	L/L	<0.3	>0.55
Hematocrit (female)	%	<30	>50	L/L	<0.3	>0.5
Leukocyte (White Blood Cell Count)	x 10^3/uL	≤2.8	≥15	x 10^9/L	≤2.8	≥15
Neutrophils	x 10^3/uL	≤1.5	No upper limit	x 10^9/L	≤1.5	No upper limit
Platelet Count	x 10^3/uL	≤75	≥700	10^9/L	≤75	≥700
Chemistry (serum or plasma)						
ALT (SGPT)	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
AST (SGOT)	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
Total Bilirubin	mg/dL	No lower limit	≥1.5 ULN	umol/L	No lower limit	≥1.5 ULN
BUN	mg/dL	No lower limit	≥30.0	mmol/L	No lower limit	≥10.71
Creatine Kinase (CK)	U/L	No lower limit	≥3 ULN	U/L	No lower limit	≥3 ULN
Sodium	mEq/L	≤125	≥155	mmol/L	≤125	≥155
Potassium	mEq/L	≤3.0	≥5.5	mmol/L	≤3.0	≥5.5
Calcium, total	mg/dL	<8.0	>11.0	mmol/L	<2.0	>2.75
Lactate Dehydrogenase (LDH)	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
Alkaline Phosphatase	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
Uric acid (male)	mg/dL	No lower limit	≥10.5	umol/L	No lower limit	≥624.75
Uric acid (female)	mg/dL	No lower limit	≥8.5	umol/L	No lower limit	≥505.75
Albumin	g/dL	≤2.6	≥6.0	g/L	≤26	≥60
Total Protein	g/dL	≤5.0	≥10.0	g/L	≤50	≥100
Chloride	mEq/L	≤85	≥120	mmol/L	≤85	≥120
Glucose (random)	mg/dL	≤45.1	≥200.0	mmol/L	≤2.48	≥11
Serum Creatinine	mg/dL	Not Applicable	>1.5 ULN	umol/L	Not Applicable	>1.5 ULN
Triglycerides	mg/dL	Not Applicable	>300	mmol/L	Not Applicable	>3.39
Gamma-Glutamyl Transferase (GGT)	U/L	Not Applicable	≥3 ULN	U/L	Not Applicable	≥3 ULN
T4, free	ng/dL	below normal range	above normal range	pmol/L	below normal range	above normal range
TSH	uU/L	below normal range	above normal range	mIU/L	below normal range	above normal range

Table 5 Criteria for Potentially Clinically Important Laboratory Values – Urinalysis

Urinalysis (qualitative dipstick)	Low PCI Criteria	High PCI Criteria
Blood (occult blood)	Not Applicable	≥ +2
Protein	Not Applicable	≥ +2
Glucose	Not Applicable	≥ +2

Clinical laboratory data will be displayed in data listings with date and study day of collection. All units will be displayed according to SI conventions for units. Out of range values will be flagged in the data listings (i.e., 'L' or 'H'). A separate listing will be provided for a subset of the chemistry, hematology, and urinalysis analytes with values classified as PCI.

The pregnancy results (positive or negative) for female subjects will be presented in a listing.

In addition, the following additional PCI criteria for liver function tests will be flagged, listed, and summarized:

- ALT (SGPT)>=2x ULN U/L
- AST (SGOT)>=2x ULN U/L
- Gamma Glutamyl Transferase>=2x ULN U/L
- Alkaline Phosphatase>=2x ULN IU/L

14.3 Vital Signs

Vital signs are assessed at Screening, Day -1 predose, and on Day 1 every 2 hours for the first 6 hours after the initial dose; vital signs then to be evaluated every 6 hours thereafter until Day 4 (EOT).

Vital signs will be summarized using descriptive statistics at Baseline and all scheduled post-Baseline visits/timepoints. The change from Baseline values will also be summarized at the scheduled post-Baseline visits/timepoints.

Vital sign values will be considered PCI if they meet the criteria listed in Table 6. The number and percentage of subjects with post-Baseline vital signs that are PCI will be summarized by treatment group at each post-Baseline visit and for overall post-Baseline. For the overall post-Baseline summaries, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI vital sign for the given parameter, visit and treatment group, and the denominator is the number of subjects with non-missing values for the given parameter, visit and treatment group. For the overall

post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI vital sign for the given parameter and treatment group, and the denominator is the number of subjects with at least 1 post-Baseline vital sign for the given parameter and treatment group. A listing of overall and of subjects with any PCI vital sign values will be provided.

Table 6 Criteria for Potentially Clinically Important Vital Signs

Vital Sign Parameter	Unit	Criteria[a]		
		Observed Value	And/Or	Change Relative to Baseline
Systolic blood pressure (supine or sitting)	mmHg	≥180	And	Increase of ≥20
		≤90	And	Decrease of ≥20
Diastolic blood pressure (supine or sitting)	mmHg	≥105	And	Increase of ≥15
		≤50	And	Decrease of ≥15
Pulse (supine or sitting)	bpm	≥120	And	Increase of ≥15
		≤50	And	Decrease of ≥15

[a] A post-baseline value is considered as a PCI value if it meets both criteria for observed value and change from baseline.

14.4 Electrocardiogram (ECG)

12-lead ECGs are collected at Screening, Day -1 predose, Days 1 through 3 (at 1 hour postdose), and EOT/ET. Observed values of ECG variables (e.g., heart rate, PR interval, QRS interval, QT interval, and QTc interval) and the changes from Day -1 (predose) at each assessment time point will be summarized by treatment group.

QTcB and QTcF will also be categorized into the following categories (msec) and the number and percentage of subjects in each category will be summarized by treatment group at each visit and for the overall post Day -1 dose maximum:

- Observed: ≤450, 451 - ≤480, 481 - ≤500, and >500; >450; >480.
- Change from Baseline: ≤10, 11 – 30, 31 – 60, and >60; >30.

For cardiologist's interpretations, the number and percentage of subjects with ECG results that are determined as normal or abnormal will be summarized at scheduled visits. The overall post-baseline worst interpretation will also be summarized (i.e. if a subject has one or more post-baseline ECG results that is/are considered as abnormal, this subject will be counted in the abnormal category). Cardiologist's interpretations will also be summarized in a shift table. The shifts from Baseline to overall post-Baseline worst interpretation will also be presented. For the by-visit shift summary, the denominator is the number of subjects with non-missing cardiologist's interpretation at Baseline and the given visit for the given treatment group. For the summaries of shift from Baseline to the overall post-Baseline worst interpretation, the denominator is the number of subjects with non-missing Baseline and at least 1 post-Baseline cardiologist's interpretation for the given treatment group.

Electrocardiogram variable values will be considered PCI if they meet the criteria listed in

Table 7. The number and percentage of subjects with post-baseline PCI values will be summarized by treatment group at each post-Baseline visit and for overall post-Baseline. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI ECG for the given parameter, visit and treatment group, and the denominator is the number of subjects with non-missing values for the given parameter, visit and treatment group. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI ECG for the given parameter and treatment group, and the denominator is the number of subjects with at least 1 post-Baseline ECG value for the given parameter and treatment group. A listing of overall and of all subjects with any PCI ECG values will be provided.

Table 7 Criteria for Potentially Clinically Important ECG Values

ECG Parameter	Unit	High PCI Criteria
QRS Interval	msec	≥120
PR Interval	msec	≥220
QTcB or QTcF	msec	>500
QTcB or QTcF: change from baseline	msec	>60

14.5 Physical Examination

Physical examinations are performed at Day -1 predose, Day 4, and at Follow-Up (surgical site only). Physical examination results (normal, abnormal, and not done) will be summarized in a frequency table by treatment group, body system and visit. A listing of overall physical examination data will be listed.

A listing related to surgical site inspection will be generated (Day 4 General Appearance, Day 7 Musculoskeletal). A shift table for surgical site inspection will be presented summarizing Day 4 and Day 7 and the shifts between the two visits (Abnormal Not Clinically Significant, Abnormal Clinically Significant).

14.6 Other Safety Variables

14.6.1 Suicidality Based on C-SSRS

The event counts and the number and percentage of subjects reporting any post Baseline suicidal ideation, suicidal behavior, or suicidality will be tabulated by treatment group. The event counts and the number and percentage of subjects reporting any post-Baseline non-suicidal self-injurious behavior will also be tabulated. For calculating the percentages, the denominator will be the number of subjects in the Safety Analysis Set.

15 CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

All summary and analyses of PK or PD of ACP-044 will be performed for the PK Analysis Set based on the actual treatment group. The treatment groups will be ACP-044 400 mg Q6H and ACP-044 1600 mg QD.

15.1 Pharmacokinetic Endpoints

The PK endpoints of the study are:

- Plasma concentrations of ACP-044 for a total of six (6) samples: Day -1 (predose), Day 1 (one sample taken between 2-3 hours and one sample taken between 4-6 hours, after the first dose of the day), Day 2 (prior to the first dose of the day), and Day 3 (one sample taken between 2-6 hours and one sample taken between 8-12 hours, after the first dose of the day)
- ACP-044 PK parameters determined using a population PK approach

Plasma concentration of ACP-044 will be summarized by treatment group using the Pharmacokinetic Analysis Set.

PK parameter summary and analysis will be discussed in a separate plan.

15.2 Pharmacokinetic/Pharmacodynamic Endpoints

The PK/PD endpoints include:

- ACP-044 exposure-response relationship:
    - for efficacy using the primary endpoint and selected secondary endpoints
    - for safety using selected TEAEs

PK/PD endpoint summary and analysis will be discussed in a separate plan.

16 INTERIM ANALYSIS

No interim analyses are planned for this study.

17 DATA MONITORING/REVIEW COMMITTEE

There is no data monitoring/review committee in this study.

18 COMPUTER METHODS

Statistical analyses will be performed using Version 9.4 (or newer) of SAS® (SAS® Institute, Inc., Cary, North Carolina) on a suitably qualified and validated environment.

Validation and quality control of the tables, listings and figures containing the results of the statistical analyses will follow appropriate standard operating procedures.

19 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

None.

20 REFERENCES

ICH E9(R1): Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials, Adopted on November 2019

FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic Guidance for Industry, Investigators, and Institutional Review Boards (March 2020)

Points to consider on implications of Coronavirus disease 4 (COVID-19) on methodological aspects of ongoing clinical 5 trials, March 25 2020.

Rubin, D. B. Inference and Missing Data. Biometrika. 1976; 63:581-92.

Centers for Disease Control and Prevention. 2018 Annual Surveillance Report of Drug-Related Risks and Outcomes — United States. Surveillance Special Report 2. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. Published August 31, 2018.

21 APPENDICES

21.1 Schedule of Events and Assessments

	Screening Period	Double-blind Treatment Period							Safety Follow-up Period	
									Clinic visit	Telephone call
Visit Day	**-28 to -2**	**Day -1 Predose**	**Day -1 (Dosing and Postdose)**	**Day 1 (Time 0)**	**Day 2**	**Day 3**	**Day 4 (EOT/ET)**	**Unscheduled Visit[m]**	**7 days after EOT**	**30 days after last dose**
Visit window (# days)		**0**	**0**	**0**	**0**	**0**	**0**		**±3**	**+4**
Informed consent	X									
Inclusion/exclusion criteria	X	X								
Medical and surgical history and demographics	X									
Physical examination	X	X					X[k]		X[k]	
Vital signs[a]	X	X		X	X	X	X	X		
Height, weight[b]	X	X								
12-lead ECG[c]	X	X		X	X	X	X			
Clinical laboratory tests	X	X					X			
Pregnancy test[d]	X	X								
Urine toxicology screen	X	X								
COVID-19 test[e]	X	X								
Clinic check in/out		X					X			
Randomization		X								
Study drug administration			X	X	X	X				
Pain assessment training[f]	X	X								
Bunionectomy			X							
Block Quality Assessment[g]			X							
11-point Numeric Rating Scale (NRS) of pain intensity[h]				X	X	X	X			
Foot and Ankle Ability Measure		X							X	
Rescue medication[i]				X	X	X	X			

21.1 Schedule of Events and Assessments (Continued)

	Screening Period	Double-blind Treatment Period							Safety Follow-up Period	
									Clinic visit	Telephone call
Visit Day	-28 to -2	Day -1 Predose	Day -1 (Dosing and Postdose)	Day 1 (Time 0)	Day 2	Day 3	Day 4 (EOT/ET)	Unscheduled Visit[m]	7 days after EOT	30 days after last dose
Visit window (# days)		**0**	**0**	**0**	**0**	**0**	**0**		**±3**	**+4**
Global evaluation of study drug				X	X	X	X			
C-SSRS	X	X					X			
PK blood sample collection[j]		X		X	X	X				
Issue diary							X[l]			
Collect diary									X	
Prior and concomitant medications	X	X	X	X	X	X	X	X	X	X
Assessment of adverse events	X	X	X	X	X	X	X	X	X	X

AE=adverse event; COVID-19=Coronavirus disease 2019; C-SSRS=Columbia-Suicide Severity Rating Scale; ECG=electrocardiogram; EOT=end of treatment; ET=early termination; SAE=serious adverse event

a Vital signs will be evaluated starting at Screening; Day -1 predose and on Day 1 every 2 hours for the first 6 hours after the initial dose. Vital signs then to be evaluated every 6 hours thereafter until Day 4/end of treatment. A 15-minute window is permitted for vital sign measurements.

b Height and BMI will only be measured and calculated at the Screening visit; weight will be measured at Screening and Day -1 predose.

c A single 12-lead ECG will be performed at Screening, on Day -1 predose, on Days 1 through 3 at 1 hour postdose (the first dose of the day), and EOT/ET. A 1-hour window is permitted for the predose ECG and a 20-minute window is permitted for the postdose ECGs.

d Applicable to women of childbearing potential. A serum pregnancy test should be performed at Screening and a urine test must be performed at clinic check-in, prior to surgery.

e COVID-19 diagnostic polymerase chain reaction (PCR) test at Screening and a rapid antigen test on Day -1 predose.

f Subjects will be trained on the process for completion of pain assessments.

g Block Quality assessments are to be performed and recorded hourly (±15 minutes) on Day -1 until the block is removed, except between 2200 and 0600 hours when the subject is asleep. Between these hours, the subject will not be awaken, except at 0300 hours (±15 minutes) when the assessment is to be completed.

h Subjects should not have ice administered to the surgical area within the 30 minutes prior to each pain assessment. Pain (efficacy) assessments are to be recorded hourly (±15 minutes) for a 12 hour interval on Day 1 beginning immediately prior to the fifth dose (time 0) administered in the study. Pain assessments will continue every 3 hours during the subsequent 12 hour interval, and then every 6 hours thereafter until 72 hours, then once daily in the morning until the follow-up clinic visit. Rescue medication use is to be recorded until the safety follow-up clinic visit. Pain intensity should also be obtained immediately prior to rescue medication being taken.

i Record name, amount, date, and time of rescue medication ingested and current pain score until the follow-up clinic visit 7 days after EOT.

j PK blood samples will be collected Day -1 (predose, within one hour of dosing), Day 1 (one sample taken between 2-3 hours and one sample taken between 4-6 hours, after the first dose of the day), Day 2 (prior to the first dose of the day), and Day 3 (one sample taken between 2-6 hours and one sample taken between 8-12 hours, after the first dose of the day); and in the event of an SAE or an AE leading to discontinuation. A 15-minute window is permitted around each nominal timepoint on Day 1 and beyond.
k The physical exam on Day 4 (EOT/ET) must include an exam of the operative site. The physical exam at the follow-up visit 7 (±3) days after the EOT procedures is for examination of the operative site only. Examination of the operative site must include visual inspection.
l Diary will be issued to take home which will be returned at the follow-up clinic visit 7 days after EOT.
m The designated study procedures are to be completed; however, the Investigator may complete any study procedure deemed necessary.

21.2 Multiple Imputation Details

The multiple imputation will be performed in a multi-step process:

1. Subjects with missing Time 0 assessments will be imputed by using the mean of the non-missing Time 0 assessments of the same treatment group.
2. Subjects that take rescue medications within 4 hours prior to a scheduled assessment will have the scheduled value replaced with the pre-rescue value for the purposes of imputation. If the scheduled value is higher than the pre-rescue value, it will not be replaced.
3. The dataset will be transposed to create a one-observation-per-patient dataset with variables representing the values at each scheduled time point (T_0 T_1 T_2 T_3 T_4 T_5 T_6 T_7 T_8 T_9 T_10 T_11 T_12 T_15 T_18 T_21 T_24 T_30 T_36 T_42 T_48 T_54 T_60 T_66 T_72). Site indicator variables (with values 0/1) will be created for the sites (with one less variable than number of sites). For the purposes of illustration, the follow text assumes 4 sites and three variables (S_1 S_2 S_3).
4. Values will be imputed assuming missing at random to create 50 complete datasets with values for all time points.
5. Subjects that drop out for adverse events or lack of efficacy will have their post-dropout missing values reset to missing.
6. Subjects with nonmissing values at 1 hour (aka T_1 variable) will have the T_1 variable replaced with max(T_0, T_1); this will include ALL subjects that do not drop out due to AE or lack of efficacy (LOE) (thus, may include values imputed in prior steps) and subjects that drop out due to AE or LOE prior to them dropping out. Proc MI will then be run including the site variables (S_1 S_2 S_3) and the two NRS variables through 1 hour: (T_0, T_1).
7. Subjects that had non-missing T_1 will have their values returned to their prior state (NOT the max of T_0, T_1), and the process will be repeated for T_2: Patients with

nonmissing values at 2 hours (aka T_2 variable) will have the T_2 variable replaced with max(T_0, T_1, T_2); this will include ALL subjects that do not drop out due to AE or LOE (thus, may include values imputed in prior steps) and subjects that drop out due to AE or LOE prior to them dropping out. Proc MI will then be run including the site variables (S_1 S_2 S_3) and the three NRS variables through 2 hours: (T_0, T_1, T_2).

8. This will be repeated iteratively up to the 72-h time point, building up to a complete dataset. This exercise yields values that are based on the distribution of the worse values through the time point being imputed for dropouts due to AE/LOE.

9. This complete dataset will be used ONLY to fill in monotone missing values starting at subject dropout. No observed values from the original dataset will be replaced and intermittent missing values and values following rescue med administration will be handled as described in above in the Section 5.2.1 when calculating AUCs. This will result in 50 records for each patient for each $AUC_{0\text{-time}}$.

The following sample code represents the imputation in step 4 above:

```
proc mi data=nrs seed=7238364 nimpute=50 out=nrs0 minimum=xx maximum=xx;
mcmc impute=full chain=multiple;
by TRT01P;
var S_1 S_2 S_3 T_0 T_1 T_2 T_3 T_4 T_5 T_6 T_7 T_8 T_9 T_10 T_11 T_12 T_15
T_18 T_21 T_24 T_30 T_36 T_42 T_48 T_54 T_60 T_66 T_72;
run;
```

The following sample code represents the imputation in step 6 above:

```
proc mi data=nrs0 seed=7238364 + seq# nimpute=1 out=nrs1 minimum=xx maximum=xx;
mcmc impute=full chain=multiple;
by imputation TRT01P;
var S_1 S_2 S_3 T_0 T_1;
run;
```

The following sample code represents the imputation in step 7 above:

```
proc mi data=nrs1 seed=7238364 + seq# nimpute=1 out=nrs2 minimum=xx maximum=xx;
mcmc impute=full chain=multiple;
by imputation TRT01P;
var S_1 S_2 S_3 T_0 T_1 T_2;
run;
```

The following list of seeds will be used when needed for MI procedures or other analyses call for a random number seed. Each will only be used once, but if additional seeds are required, they will be chosen by adding one to a digit from the below list in an organized manner. For example, the primary analysis may use the list in order as it appears; a sensitivity may add one to the final digit of each; a second sensitivity may add one to the first digit of each. The actual seeds used will be captured in the analysis dataset programing specifications.

Seed 7238364 will be used for the multiple imputation other than placebo-based approach.

Seed 1592743 will be used for the placebo based multiple imputation.

Replace Steps 6-8 with the following sample code to achieve control-based copy-placebo imputation for missing data due to AE and LOE:

```
PROC MI DATA=MONO SEED=xxxxx NIMPUTE=1 OUT=OUTM1 minimum=xx maximum=xx;
CLASS TRT01PN;
By _IMPUTATION_;
VAR S_1 S_2 S_3 AVAL1 AVAL2 AVAL3……;
MONOTONE REG (AVAL2 AVAL3……);
MNAR MODEL (AVAL1 AVAL2 AVAL3…../ MODELOBS = (TRT01PN = '1'));
RUN;
```

21.3 Summary of Version Changes

Version No:	Document History Description of Update	Author(s)	Version Date
1.0	Original Version	[redacted]	14 March 2022