Document Type:	Statistical Analysis Plan
Official Title:	A Pilot Study Assessing the Treatment Responsiveness of a Novel Osteoarthritis Stiffness Scale
NCT Number:	NCT03570554
Document Date:	25-SEP-2019

Statistical Analysis Plan Approval Form

Clinical Protocol Title: A Pilot Study Assessing the Treatment Responsiveness of a Novel Osteoarthritis Stiffness Scale	
Investigational Product: BAY 117031 / Naproxen Sodium	
Trial IMPACT No.: 19783	**Phase: 2**
Version No./Date*(dd/mmm/yyyy)*: V 1.0 Final / 25Sep2019	

Signing this form constitutes agreement with the Statistical Analysis Plan or any Amendment.

APPROVED BY:

Role	Name	Signature	Date
Clinical Development Services Representative	PPD	PPD	25-Sep-2019
Trial Statistician	PPD		9/25/19
Medical Affairs Representative	PPD		25 Sept 2019
CRO Statistician CCI	PPD		

Bayer HealthCare

19783

A Pilot Study Assessing the Treatment Responsiveness of a Novel Osteoarthritis Stiffness Scale

25SEP2019

Statistical Analysis Plan

Version *1.0*

Prepared by:

CCI [redacted]

TABLE OF CONTENTS

List of Abbreviations

AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BASS	Brief Arthritis Stiffness Scale
BMI	Body Mass Index
CABG	Coronary Artery Bypass Graft Surgery
CI	Confidence Interval
COX	Cyclooxygenase
CRF	Case Report Form
CSR	Clinical Study Report
DSMB	data safety monitoring board
ER	Extended Release
eCRF	(electronic) Case Report Form
FDA	Food and Drug Administration
ICF	Informed Consent Form
IMP	Investigational Medicinal Product
ITT	Intent-to-Treat
IUD	Intra-uterine Device
MedDRA	Medical Dictionary for Regulatory Activities
NRS	Numerical Rating Scale
NSAID	Nonsteroidal Anti-inflammatory Drug
OA	Osteoarthritis
OTC	Over-the-Counter
PI	Pain Intensity
PID	Pain Intensity Differences
PP	Per Protocol
PT	Preferred Term
PVG	Pharmacovigilance
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SOC	System Organ Class
SPI	Sum of Pain Intensity
SPID	Summed Pain Intensity Difference
STEPP	Staircase-Evoked Pain Procedure
SUSAR	Suspected Unexpected Serious Adverse Reaction
TEAE	Treatment-Emergent Adverse Events
WHO	World Health Organization
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

1. Introduction

Osteoarthritis (OA) is a common, chronic, progressive, skeletal, degenerative disorder that frequently affects several joints such as knee, hip, spine and hands. Osteoarthritis is associated with damage to the cartilage and surrounding tissues and characterized by pain, stiffness, and loss of function.(1) Being a highly prevalent disease and a main form of arthritis, OA is the most common cause of disability in the U.S. In fact, 52.5 million U.S. adults (approximately 1 in 4) aged ≥18 years have self-reported having doctor-diagnosed arthritis according to 2010-2012 data collected by the National Health Interview Survey. (2) The incidence of large joint OA in the U.S., particularly of the knee, is predicted to increase with the aging of the population and obesity epidemic. (3)

Currently, a variety of prescription and over-the-counter (OTC) medications are available. Acetaminophen has been a mainstay of OA pharmacological management for many years and has been tested in many clinical trials although its effectiveness in clinical trials appears limited. (4) An extended release (ER) formulation containing acetaminophen 650 mg in a bilayer form has been shown to be more efficacious than placebo in hip and knee OA. (5) Each ER caplet contains 650 mg of acetaminophen. The maximum daily dose is 3900 mg.

Aleve®, naproxen sodium, the sodium salt of naproxen, was developed as an analgesic because it is more rapidly absorbed than naproxen. It is available as an OTC analgesic/antipyretic that temporarily reduces fever and relieves minor aches and pains due to: minor pain of arthritis, muscular aches, backache, menstrual cramps, headache, toothache, and the common cold. OTC doses of naproxen sodium have been shown to be efficacious for relieving OA pain. (6, 7, 8) Each tablet contains 220 mg of naproxen sodium with a maximum daily dose of 660 mg.

Celecoxib, a COX-2 selective NSAID, is a commonly used prescription analgesic for OA and its efficacy is supported by a large body of data. (9) For relief of the signs and symptoms of OA the recommended oral dose is 200 mg per day.

Along with pain, stiffness is a frequent symptom of OA. (6) The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), a commonly used subject reported outcome (PRO) instrument for the measurement of OA-related joint pain and physical function includes a 2-item subscale for measuring joint stiffness. (10) However, the WOMAC's 2-item subscale for measuring joint stiffness shows low test-retest reliability. (11) FDA has informed Bayer that the WOMAC stiffness subscale has not been accepted as a stand-alone measure suitable for use as an endpoint.

In order to examine the efficacy of treatments on improving joint stiffness, the Brief Arthritis Stiffness Scale (BASS), a PRO instrument designed specifically to measure stiffness was developed and validated for use in clinical trials. A non-interventional study has previously been conducted to examine the psychometric properties of the BASS. The study provided evidence for the reliability, construct validity, and discriminant validity of both the daily and weekly average BASS score, as well as preliminary responder thresholds. Further evaluation of the BASS to assess the instrument's responsiveness to treatment is needed. A study including an

active intervention will provide information to formally evaluate the instrument's ability to detect change and provide utility for the responder definition.

This placebo-controlled clinical trial will assess the effects of naproxen sodium, acetaminophen ER and celecoxib on stiffness in subjects with OA. Furthermore, data will be collected to confirm the psychometric evidence supporting the use of the BASS as a subject reported outcome in assessing the effect of treatment in OA.

2. Objectives

Primary

- To assess the responsiveness of the Brief Arthritis Stiffness Scale (BASS) for detecting treatment effects of common OTC analgesics and a common prescription analgesic in subjects with knee OA.

Exploratory

- To assess the assay sensitivity of the **ST**aircase-**E**voked **P**ain **P**rocedure (STEPP™) to detect treatment effects of commonly used analgesics compared to placebo and to detect differences between active treatments;

- To determine/identify effect size measurements, determine single dose vs. multi-dose efficacy, identify optimal sample size, and logistical issues for planning comparative analgesic efficacy studies using the STEPP.

- To assess levels of activity captured by the accelerometer between active treatments and placebo during Days 1 (partial), 2 and 3 of the treatment phase.

- To evaluate the safety and tolerability of naproxen sodium, celecoxib and acetaminophen ER.

3. Investigational Plan

3.1. Overall Study Design and Plan

This is a multi-center, randomized, double-blind, four period crossover, placebo-controlled pilot study in adult male and female subjects between 40 and 80 years of age with pain due to osteoarthritis of the knee. At the completion of the Screening phase and baseline, eligible subjects will be randomized in a 1:1:1:1 crossover fashion (A:B:D:C, B:C:A:D, C:D:B:A, D:A:C:B) to a treatment sequence administered as shown in Table 1.:

Table 1: Treatments administered

	Dose Time	Treatment A Naproxen	Treatment B Acetaminophen ER	Treatment C Celecoxib	Treatment D Placebo
Day 1 Visit 2	8 AM	440 mg	1300 mg	100 mg	Placebo
	4 PM	Placebo	1300 mg	Placebo	Placebo
	8 PM	220 mg	Placebo	100 mg	Placebo
	12 AM	Placebo	1300 mg	Placebo	Placebo
Day 2 At home	8 AM	440 mg	1300 mg	100 mg	Placebo
	4 PM	Placebo	1300 mg	Placebo	Placebo
	8 PM	220 mg	Placebo	100 mg	Placebo
	12 AM	Placebo	1300 mg	Placebo	Placebo
Day 3 At home	8 AM	440 mg	1300 mg	100 mg	Placebo
	4 PM	Placebo	1300 mg	Placebo	Placebo
	8 PM	220 mg	Placebo	100 mg	Placebo
	12 AM	Placebo	1300 mg	Placebo	Placebo
Day 4 Visit 3	8 AM	440 mg	1300 mg	100 mg	Placebo

The study consists of up to 14-day Screening Phase, a 4(+3) day Screening Washout Phase, and a Treatment Phase consisting of four 4-day treatment periods separated by washout periods lasting several days. The study center may schedule screening testing on multiple days as needed, provided all screening tests are done within 14 days. The target knee will be established at the screening visit. Approximately 192 subjects will be screened with 62 eligible to start the Screening Washout Phase. Approximately 12 subjects are anticipated to fail baseline randomization criteria to obtain approximately 50 randomized subjects of whom approximately 10 are anticipated to drop out leaving approximately 40 completed subjects. Duration of trial participation will be approximately 68 days as detailed in Figure 1. See Section **Error! Reference source not found.**for complete schedule of events.

Note that concomitant pain therapies, including topicals, and supplements must be discontinued at the start of the Screening Washout Phase (Day -4 to -1 +3 days) through trial completion. Subjects can take acetaminophen ER(up to 2000 mg/day) provided by the study site during this Screening Washout Phase, but not within 24 hours of the Randomization visit (Day 1 of Treatment Period 1). No other analgesic medications are permitted.

Subjects will return to the study center at the beginning of each treatment period, upon completion of the Screening Washout Phase or between-treatment washout period. In the original version of the protocol, in order to be eligible for randomization prior to Treatment Period 1, subjects needed to meet the following randomization criteria: self-reported average score of $\geq$5 on the WOMAC pain subscale (0-10 NRS, 24-hour recall version) related to OA pain in the target knee; current pain intensity score of $\geq$4 and $\leq$9 on the 0-10 NRS; and, a score of $\geq$5 on the 0-10 NRS for average pain over the past 24-hours. In addition, the Day 1 baseline post-STEPP current pain intensity 0-10 NRS score must be $\geq$6, and $\geq$1 unit higher than the pre-STEPP current pain intensity 0-10 NRS score, and the subject must experience knee stiffness at

least 5-7 days per week over the past month as reported on the Baseline Stiffness Questionnaire at the Screening Visit.

In Protocol Version 3.0, the randomization criteria was amended: 48-hour WOMAC pain subscale average score of ≥3 and/or a 24-hour Average Pain Intensity score ≥3 on the 0-10 NRS at Screening Visit, 48-hour WOMAC pain subscale average score of≥4 on the 0-10 NRS at Visit 2, Treatment Period 1 of 4, 24-hour Average Pain Intensity score ≥3 at Visit 2, Treatment Period 1 of 4, current pain intensity score of ≥ 4 and ≤9 on the 0-10 NRS at Visit 2, Treatment Period 1 of 4 and Day 1 baseline post-STEPP current pain intensity score ≥5, and ≥1 unit higher than the pre-STEPP current pain intensity 0-10 NRS score, and the subject must experience knee stiffness at least 5-7 days per week over the past month as reported on the Baseline Stiffness Questionnaire at the Screening Visit.

Subjects who do not meet all the randomization criteria will be considered a Screening failure and will not be randomized into the study. Subjects who qualify will be asked additional questions related to knee pain, stiffness and function.

Randomized subjects will receive their treatment period specific investigational medicinal product (IMP) or treatment kit that includes study medication at the start of each treatment period. Randomization occurs only once during the start of the first treatment period.

Withdrawn subjects may be classified as a screening failure, an inadequate pain reporter, or a dropout, depending on the time point of withdrawal. Withdrawn subjects will not be replaced in this study.

During the completion of Visit 2 (Day 1), subjects will be given a diary to record daily BASS scores prior to next visit and fitted with an accelerometer to measure the number of steps taken and level of activity (e.g., low, medium, high). Accelerometer data collection will start after completion of the Day 1 activities prior to leaving the study site, however only data from full days (Day 2 and Day 3) will be used in data analysis

Subjects will start a new Washout Period following the completion of Visit 3 (Day 4) of each Treatment Period (except Treatment Period 4). The Washout Period will last from 3 to 7 days after Visit 3. Subjects will not be allowed to take any pain medication not dispensed by the study center. Subjects can take acetaminophen ER (up to 2000 mg/day) as a rescue medication provided by the study site during this Washout Period but not within 24 hours of Day 1 of the next Treatment Period. Subjects will be scheduled to return to the study center to start the next Treatment Period (X of 4) for Day 1, Visit 2.

Figure 1: Design Overview

	Screening Phase	Screening Washout	Treatment Phase (4 arms)					Washout Phase	Follow up Phase
Trial days	Screening Visit 1 Day -18 to -5	Day -4 to -1 (+3 d)	Start Treatment Period Visit 2 Day 1			Continue Treatment Day 2 & 3	End Treatment Period Visit 3 Day 4	Between Treatment Periods	End of Study
	Accurate Pain Reporting Training Clinical Trial Expectations Training Baseline Stiffness Questionnaire WOMAC Pain WOMAC Stiffness WOMAC Function Stiffness Severity Question 24-hour average pain (0-10 NRS) STEPP test Regional Pain Scale	Washout before 1st IMP dose	Daily BASS WOMAC Pain ≥4 24-hour average pain ≥4 (0-10 NRS) Current pain ≥4 and ≤9 (0-10 NRS) Post STEPP pain ≥5 and ≥1 unit higher than pre-STEPP pain (0-10 NRS)	*	WOMAC Stiffness WOMAC Function Daily BASS 1st IMP dose   2nd, 3rd & 4th IMP dose Accelero-meter use after completion of study procedures	Daily BASS prior to 1st IMP dose IMP doses 2, 3 & 4 each day Accelerometer use	Daily BASS WOMAC Pain WOMAC Stiffness WOMAC Function Morning IMP dose   PGIC Treatment Preference[a]	3 to 7 days of washout needed before starting the next treatment in the sequence After washout, repeat Treatment Phase for the remaining treatment(s) in the sequence (X of 4)	7-10 days after completion of treatment period 4 of 4 Phone call Qualitative Interview[a]

* = Randomized to a 4 arm blinded treatment sequence (only for treatment 1 of 4)

 = Staircase-Evoked Pain Procedure (STEPP) completed pre-dose, and then 2, 4, 6 and 8 hours post-dose

 = 0-10 point current pain NRS at pre-dose immediately before and 2-5 minutes after the STEPP, which is also completed at 2, 4, 6 and 8 hours post-dose. Additionally, hourly current pain intensity NRS at 1, 3, 5, and 7 hours post-dose when not completing the STEPP

[a]Treatment period 4 of 4 only or if the subject is withdrawn from the study prior to completion

3.2. Study Endpoints

3.2.1. Primary Efficacy Endpoint

The primary efficacy endpoint is:

- Sum of change from baseline (Day 1 of the Treatment Period) in BASS score over the 4 day treatment period

3.2.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints include

- Absolute BASS score at each time point.
- Day 4 change from baseline (Day 1 of the treatment period) in BASS score

Note: Key pairwise comparisons of absolute BASS score and change from baseline BASS score will be assessed, including, but not limited to: naproxen versus placebo, celecoxib versus placebo, acetaminophen ER versus placebo, naproxen versus acetaminophen ER and naproxen versus celecoxib.

3.2.3. Exploratory Efficacy Endpoints

- Day 1 and Day 4 post-STEPP sum of pain intensity difference (SPID) at 4, 6, and 8 hours post-dose, based on 0-10 NRS;
- Day 1 and Day 4 post-STEPP sum of pain intensity (SPI) at 4, 6, and 8 hours post-dose, based on 0-10 NRS;
- Day 1 and Day 4 post-STEPP pain intensity difference (PID) from baseline at 2, 4, 6, and 8 hours post-dose, based on 0-10 NRS;
- Day 1 and Day 4 post-STEPP pain intensity (PI) at 2, 4, 6, and 8 hours post-dose, based on 0-10 NRS;
- Current hourly PI using the 0-10 NRS pre-dose, and from 1 to 8 hours post-dose;
- Mean change from baseline (Day 1 of the Treatment Period) to end of Treatment Period on the Western Ontario and McMaster Universities Arthritis (WOMAC) pain subscale, 48-hour recall version, assessed on Day 1 and Day 4 including an assessment of all key pairwise comparisons;
- Mean change from baseline to end of Treatment Period on WOMAC stiffness subscale, 48-hour recall version, assessed on Day 1 and Day 4 including an assessment of all key pairwise comparisons;
- Absolute WOMAC stiffness subscale scores, 48-hour recall version, for each study treatment at each time point;
- Physical activity level outcomes during Days 1 (partial), 2 and 3 as measured by the accelerometer;

- Mean change from baseline to end of treatment period on WOMAC function subscale, 48-hour recall version, assessed on Day 1 and Day 4, including an assessment of all key pairwise comparisons;
- Patient Global Impression of Change (PGIC) on Day 4 of each Treatment Period, including PGIC for OA Pain (PGIC-P) and PGIC for OA Symptoms (PGIC-S);
- Subject preference for treatment assessed at the end of the study;
- Qualitative (semi-structured) interview to assess subjects' experience with treatments, and experience with the STEPP, relevance to their disease status, how to improve it, etc.
- Level of activities during Days 1 (partial), 2 and 3 as measured by the accelerometer (Modus StepWatch4).
    - Number of steps per day
    - Minutes of active vs. inactive per day
    - Percent of time in low, medium or high activity
    - Peak performance of the day (30 minutes)
    - Maximum performance (1, 5, 20, 60 minutes)
    - Cadence

3.2.4. Safety Endpoints

The following laboratory tests will be performed at screening:

- Hematology (Hemoglobin, Hematocrit, RBC count, Platelet count, WBC count)
- Serum Chemistry (BUN, Creatinine, Glucose, CA++, NA+, K+, Cl-, Total CO2 (Bicarbonate), AST, ALT, Total Bilirubin, Alkaline Phosphatase, Albumin, Total protein)
- Urinalysis (Specific gravity, pH, Glucose (qual), Protein (qual), Blood (qual), Ketones, Leukocyte Esterase, Nitrite, Bilirubin, Urobilinogen)

Adverse events will be collected throughout the treatment and safety follow-up periods and will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) dictionary, version 21.1. Only treatment-emergent AEs will be included, i.e., AEs that begin or worsen after the first dose of the investigational products in the Treatment Phase. The number and percent of subjects who experience any event, by System Organ Class (SOC), and by Preferred Term will be displayed by treatment group. Tables will also be produced by intensity and relationship to investigational product. Seriousness, intensity, relationship to investigational product, duration, and outcome will also be listed.

3.3. Treatments

Subjects will follow the dosing schedule as outlined in Table 1. The blister package will be blinded to the specific treatments in the sequence. All subjects will randomly receive one drug treatment (A, B, C, or D) during each dosing period. Depending upon which treatment in the sequence, they could be taken naproxen sodium, acetaminophen ER, celecoxib or placebo.

4. General Statistical Considerations

All statistical analyses will be conducted using SAS Version 9.3 or higher (SAS Institute, Cary, North Carolina).

CCI [redacted]

Continuous data will be summarized by treatment group using descriptive statistics (n, mean, median, standard deviation [SD], minimum, and maximum). Categorical data will be summarized by treatment group using frequency tables (frequencies and percentages). Mean and median will be presented to one decimal place beyond the precision with which the data was captured. SD will be presented to two decimal places beyond the precision with which the data was captured. Minimum and maximum will be presented to the precision with which the data was captured. When count data are presented, the percentage will be suppressed when the count is zero in order to draw attention to the non-zero counts. A row denoted "Missing" will be included in count tabulations where specified on the shells to account for dropouts and missing values. The denominator for all percentages will be the number of subjects in that treatment group within the analysis set of interest, unless otherwise specified.

All study-related raw data that support the corresponding tables and figures will be presented in data listings. Additional data listings may be generated as needed. All table, listing and figure (TLF) shells will appear in landscape format employing Courier New 9 point font. Unless otherwise noted, all tables will summarize subject results by treatment group sorted in the following order: Naproxen Sodium, Acetaminophen ER, Celecoxib, Placebo. As this is a crossover study, we expect that individual subjects will be included in all treatment groups. Unless otherwise noted, all data listings will be sorted by treatment group and subject identification number defined as investigator identification number concatenated with the subject number.

The baseline value for an assessment will be defined as the last non-missing measurement including unscheduled assessments before the first dose of IMP for each treatment period. Change from baseline is defined as the post-baseline value minus the baseline value for the given assessment.

4.1. Sample Size

Because there are no reference data on the impact of naproxen, acetaminophen ER and placebo on stiffness with the utilization of the BASS scale, an estimated 50 subjects randomized to achieve approximately 40 subjects completing the study are considered adequate for this pilot study. Screening will continue until 50 subjects are randomized in the study.

4.2. Randomization and Blinding

At the beginning of the first treatment period, after completion of the pre-treatment baseline procedures/assessments, subjects who meet the entry criteria will be sequentially assigned to a unique number in ascending order (randomization number, RNR) according to the randomization schedule prepared prior to the study.

Subjects will be numbered according to the following scheme:

14001XXXX

Whereas the "Xs" will be replaced with a four digit sequentially assigned number as each subject enters the study (e.g., first subject number will be PPD).

Once a number has been assigned to a subject, it cannot be reassigned to another subject.

Subjects who do not meet all of the randomization criteria will not be randomized. Upon successful completion of all remaining inclusion and exclusion criteria, subjects enter the Treatment Phase and are randomized into one of four blinded treatment sequences using a Williams design to ensure a balanced stratification. See example below; treatments are defined in Table 1:

A B D C
B C A D
C D B A
D A C B

Subjects enrolled in the study, investigators and their staff involved in protocol procedures or data collection analysis will be blinded to the identity of the treatment sequence. The study monitor will conduct product accountability during and after database lock. To preserve the blinding, all investigational products will be over encapsulated and prepackaged according to the randomization schedule.

4.3. Analysis Sets

Three populations will be identified in this study. Subjects' inclusion status in each of the analysis sets will be presented in a data listing.

4.3.1. Intent-to-Treat (ITT)

The ITT will consist of all subjects who were randomized and have taken at least one dose of IMP. ITT population will be used for all efficacy analyses to assess the robustness of the results for all primary, secondary, and exploratory efficacy endpoints. All analyses using the ITT will group participants according to planned treatment.

4.3.2. Per Protocol (PP)

The PP will consist of all subjects in ITT who do not have major protocol violations that could affect the evaluability of the primary efficacy parameter. The PP population will be used as the primary analysis for all efficacy parameters. Protocol deviations, including assessment of significance, will be identified prior to database lock and study unblinding. All analyses using the PP will group participants according to treatment sequence actually received. The following incidences will be considered as a protocol violation and excluded from the PP analysis if any one of the following occurred during the treatment period:

- Patient took rescue medication day before Day 1of Treatment Period or Day 1 was outside allowable visit window;
- Patient took a prohibited concomitant medication preceding or during the study Treatment Period;
- Patient failed to continue to meet the inclusion/exclusion criteria prior to the Treatment Period;
- Patient compliance rate for that treatment period is below 75%;
- Patient did not complete required assessments for the Treatment Period

Subjects may be invalid for one period but valid for another period. Therefore, invalidity for one period does not make them invalid for all. Protocol violations which lead to exclusion from the per-protocol group will be summarized and listed.

4.3.3. Safety

The Safety set is comprised of all subjects who received at least one dose of IMP. The Safety set is usually used to tabulate all of the safety information for a study, such as adverse events, treatment compliance, laboratory results, and vital signs.

4.4. Assessment Windows

4.4.1. Treatment Periods

For classifying AEs, the treatment periods will be defined as:

- Treatment Period 1: The onset date of the AE is on or after the start date of study medication in treatment period 1 but before the start date of the study medication in treatment period 2.
- Treatment Period 2: The onset date of the AE is on or after the start date of study medication in treatment period 2 but before the start date of the study medication in treatment period 3.
- Treatment Period 3: The onset date of the AE is on or after the start date of study medication in treatment period 3 but before the start date of the study medication in treatment period 4.
- Treatment Period4: The onset date of the AE is on or after the start date of study medication in treatment period 4.

Partial dates will be imputed for AE start dates in order to determine the treatment period to which the event will be assigned.

- If the onset date is completely missing, onset date is set to the date of first dose and will be considered TEAE for all treatment periods.
- If the year is present and the month is missing, then the month is set to January. If the year is same as the year of the date of first dose and the AE end date is not prior to the date of first dose, then onset date is set to the date of first dose.
- If the month and year are present and the day is missing, then the day is set to the 1st day of month. If the month and year are same as the month and year of the date of first dose and the AE end date is not prior to the date of first dose, then onset date is set to the date of first dose.
- If the AE end date is present, then the imputed start date will be no later than the end date.

Partial dates will be imputed for medication start dates in order to determine if a medication is prior or concomitant.

- Partial start dates of prior and concomitant medications will be assumed to be the earliest possible date consistent with the partial date.
- If the year is present and the month is missing, then the month is set to January.
- If the month and year are present and the day is missing, then the day is set to the 1st day of month.
- If the year is missing, then the year will be assumed to be the year part of the subject's informed consent date.

Partial or missing medication stop dates for medications that are not ongoing will be handled as follows:

- Partial stop dates of prior and concomitant medications will be assumed to be the latest possible date consistent with the partial date.

- If the year is present and the month is missing, then the month is set to December.
- If the month and year are present and the day is missing, then the day is set to the last day of month.
- For subjects who are treated, if the year is missing and the month or day are not missing, then the year will be assumed to be the year part of the subject's last recorded study visit date.
- For subjects who are not treated, if the year is missing and the month or day is not missing, then the year will be assumed to be the year part of the subject's discontinuation date.
- If the complete stop date is missing, the stop date will be considered to be either the subject's last recorded visit date for treated subjects or the subject's discontinuation date for non-treated subjects.

5. Subject Disposition

5.1. Study Sets

The number of subjects who were enrolled and the number of subjects within each analysis set (ITT, PP, and Safety) will be summarized for each treatment sequence and treatment period for all randomized subjects. A by-subject listing indicating the subject's inclusion in each analysis set and reason(s) that the subject is excluded from an analysis set will be presented.

5.2. Disposition

The counts and percentages of subjects who complete or discontinue from the study treatment will be presented based on the number of subjects in each treatment sequence and overall for all randomized subjects. The count and percentage of subjects who complete or discontinue from the study will be summarized in a similar manner. Reasons for discontinuation of treatment for subjects who do not complete the study treatment and reasons for not completing the study will be summarized for each treatment sequence as Treatment Disposition and Subject Disposition, respectively. All percentages will be based on the number of subjects randomized. Subject disposition data will be listed as well.

5.3. Protocol Deviations

Protocol deviations will be tracked by the clinical team on an on-going basis. Specific criteria for what constitutes a significant protocol deviation will be determined by the clinical team. A blinded review of the deviation log collected by the clinical group will be conducted prior to database lock. All deviations will be listed.

Protocol violations, which disqualify a subject from the per-protocol set, will also be tracked on an on-going basis. These violations will be categorized and summarized by treatment sequence within each treatment period and listed by subject.

6. Demographics and Baseline Characteristics

6.1. Demographics

Demographic information and baseline characteristics collected at screening will be summarized for the ITT, PP, and Safety populations. Continuous variables, including age (years), baseline weight (kg), baseline height (cm), and baseline body mass index (BMI) (kg/m^2) will be summarized using descriptive statistics for each treatment sequence and overall. Categorical variables including sex (Male, Female), ethnicity (Hispanic or Latino, Not Hispanic or Latino), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Multiple, and Other), and native language (English, Spanish, Chinese, Tagalog, Vietnamese, Arabic, French, Korean, Russian, German, Other) will be summarized by reporting the number and percentage of subjects in each category for each treatment sequence and overall. All demographic and baseline characteristics will be listed for all subjects in the Safety set.

Age will be calculated as (date of Screening visit – date of birth)/365.25.

BMI is calculated as (body weight in kilograms) / (height in meters)^2.

In addition, baseline (pre-dose) BASS score and post-STEPP NRS for each treatment period will be summarized using descriptive statistics for each treatment sequence and overall.

6.2. Alcohol, Tobacco and Drug Usage

Descriptive statistics will be provided by treatment sequence for subject alcohol history (Yes, No), tobacco usage (Current, Previous, Never), and whether illegal drug use history was discussed with the subject (Yes, No). Percentages will be based on the total number of subjects in the ITT, and results will be listed by subject.

6.3. Medical History

6.3.1. General Medical History

Medical history will be collected in the Electronic Case Report Form (eCRF) at the screening and baseline visits and verbatim terms will be coded classified on the basis of Medical Dictionary for Regulatory Activities (MedDRA) terminology, which will be updated whenever possible throughout the life of the study. Medical history will be summarized by the number and percentage of subjects with any medical history reported by coded system organ class, preferred term, and treatment sequence for the Safety population. A subject listing will also be included with start date and end date or ongoing status.

6.4. Inclusion and Exclusion Criteria

Inclusion and exclusion criteria deviations noted in the eCRF will be presented for the all subjects in a data listing.

7. Treatments and Medications

7.1. Prior and Concomitant Medications

Any medicinal product, prescribed or over-the-counter (OTC), including herbal products, vitamins, and minerals is considered a concomitant medication. Any medications used during the study will be coded with the World Health Organization Drug Dictionary (WHODD), version dated September 2018.

Prior/concomitant medication use will be recorded in the eCRF beginning at Screening visit before the first dose of study medication and will be recorded at the subsequent visits as well. Any changes in prior or concomitant medications will also be recorded in the subject's eCRF.

Summary of prior and concomitant medications by treatment group and preferred term will be provided. All prior and concomitant medications will be presented in data listings. Prior and concomitant medications will be summarized for the safety set. Rescue medications will be flagged in the listing.

7.2. Study Treatments

Subjects will follow the dosing schedule as outlined in Table 1. The blister package will be blinded to the specific treatments in the sequence. All subjects will randomly receive one drug treatment (A, B, C or D) during each dosing period. Depending upon which treatment in the sequence, they could be taking naproxen sodium, acetaminophen ER, celecoxib or placebo. Subjects will administer the drug with a full glass of non-refrigerated, non-carbonated water (about 8 ounces or 240 mL).

7.2.1. Extent of Exposure and Compliance

The distribution of the study medication will be supervised by a member of the Investigator's team. The subject will bring their dosing blister package (Med-ic®) with them to each visit. The Med-ic blister electronically records the date and time when the subject depresses the blister to take a dose of IMP. The study center will download and review the dosing data from the Med-ic blister package for compliance. Any deviations related to dosing outside of a $\pm$ 1 hour window will be documented in the subject's medical record. A dose will be considered compliant if it is taken within a $\pm$ 2 hour window of the expected time. Total exposure and percent compliance will be summarized for the Safety and ITT sets, and a by-subject listing will be included on the Safety set. Subject use of rescue medication will also be summarized along with IMP and included in a listing.

8. Efficacy Analysis

Efficacy data will be analyzed utilizing the PP set, with ITT set used as a sensitivity analysis. Data imputation may occur if there are any missing BASS scores during the study. They will be imputed using linear interpolation (trapezoid rule – for between periods) or extrapolation (last observation carried forward within treatment period) approach depending on availability of data surrounding the missing value. For instances where a patient only partially answers some questions in the WOMAC questionnaire (left some answers blank) the standard WOMAC method (12) will be used, where the mean of non-missing sub-scale items is imputed for the missing ones when no more than 1 stiffness question, 2 pain questions, and 4 physical function questions are missing. The following key pairwise comparisons may be tested for endpoints as described in sections below: naproxen versus placebo, celecoxib versus placebo, acetaminophen versus placebo, naproxen versus acetaminophen and naproxen versus celecoxib.

8.1. Primary Efficacy Endpoint

The daily BASS is designed to measure the severity of any stiffness experienced in the target knee over a 24-hour period. Subjects who had stiffness over the 24-hour period, are asked four questions on a 0-10 scale, with 0 representing no stiffness and 10 representing worst stiffness imaginable:

CCI

[redacted]

The primary efficacy endpoint is the sum of change from baseline (Day 1 of the Treatment Period) in BASS over the 4-day treatment period. It will be calculated by summing the values of all four questions scale on the BASS questionnaire and summing the change-from-baseline calculations at each post-baseline time point.

$$(1)\ \sum_{i=1}^{3} f_{i+1} - f_1$$

Where f_1 is BASS score (sum of four questions, 0-40 scale) at the baseline (Day 1) time point, and ($f_{i+1} - f_1$) represents the change in baseline from BASS score at each day post-baseline. The subjects complete the BASS on Day 1 prior to first dose in the treatment period, and at Day 2, Day 3, and Day 4 in each treatment period.

8.1.1. Primary Analysis

CCI

[redacted]

CCI

[redacted]

8.2. Secondary Efficacy Endpoint

8.2.1. Absolute BASS Score at Each Time Point

CCI

[redacted]

8.2.2. Day 4 Change from Baseline in BASS Score

CCI

[redacted]

8.3. Exploratory Efficacy Endpoints

8.3.1. Day 1 and Day 4 Post-STEPP Sum of Pain Intensity Difference (SPID) at 4, 6, and 8 Hours Post-Dose

On Days 1 and 4 of each treatment period after dosing, subjects will undergo the STEPP procedure at 2, 4, 6, and 8 hours. The NRS is a numeric scale from 0-10 (0= no pain and 10=extreme pain) which measures pain at Baseline (pre-Dose), both before and after the STEPP procedure, as well as after the STEPP procedure at 2, 4, 6, and 8 hours.

Pain intensity differences (PID) will be derived by subtracting the pain intensity (using NRS 0-10, 0=no pain and 10=extreme pain) at the baseline (post-STEPP) time point from the post dose intensity score (baseline score – post-baseline score). A positive difference is indicative of improvement. Summed Pain Intensity Difference will be calculated as follows:

$$(2)\ SPID = \sum_{i}^{n} (f_0 - f_i)(t_i - t_{i-1})$$

Where f_i is the NRS score at i hours after dose of IMP, f_0 represents the baseline post-STEPP NRS score, and $(t_i - t_{i-1})$ represents the duration in hours since the last time point, with t_0 representing the baseline time point. The post-baseline time points will be at 2, 4, 6, and 8 hours. Sum of Pain Intensity Difference will be calculated at n=4, 6, and 8 hours and summarized descriptively by period and total in a table. At 8 hours on Day 1 and Day 4, the key pairwise comparisons will be performed without multiplicity adjustment, with confidence intervals and nominal p-values reported.

8.3.2. Day 1 and Day 4 Post-STEPP Sum of Pain Intensity (SPI) at 4, 6, and 8 Hours Post-Dose

Sum of pain intensity will be derived by summing the NRS scores at each post-dose time point multiplied by the interval between measurements. Sum of Pain Intensity will be calculated as follows:

$$(3)\ SPI = \sum_{i}^{n} (f_i)(t_i - t_{i-1})$$

Where f_i is the NRS score at i hours after dose of IMP, f_0 represents the baseline post-STEPP NRS score, and $(t_i - t_{i-1})$ represents the duration in hours since the last time point, with t_0 representing the baseline time point. The post-baseline time points will be at 2, 4, 6, and 8 hours. Sum of Pain Intensity will be calculated at n=4, 6, and 8 hours and summarized descriptively in a table.

8.3.3. Day 1 and Day 4 Post-STEPP Pain Intensity Difference (PID) at 2, 4, 6, and 8 Hours Post-Dose

Pain intensity differences will be calculated as indicated in **Error! Reference source not found.**, and measured at 2, 4, 6, and 8 hours post-dose. The results will be displayed in a summary table by visit, plotted and listed by subject.

8.3.4. Day 1 and Day 4 Post-STEPP Pain Intensity (PI) at 2, 4, 6, and 8 Hours Post-Dose

The pain intensity scores based on the 0-10 NRS will be presented at 2, 4, 6, and 8 hours post-dose, and the results will be displayed in a summary table by visit and listed by subject.

8.3.5. Current Hourly PI Using the 0-10 NRS Pre-Dose, and from 1 to 8 Hours Post-Dose

The pain intensity scores based on 0-10 NRS will be measured every hour for eight hours following first dose on Days 1 and 4. We will summarize all NRS scores on Days 1 and 4 by time point in a summary table, figure and list by subject.

8.3.6. Mean Change from Baseline to End of Treatment Period on the WOMAC Pain Subscale, 48-Hour Recall Version

The WOMAC consists of 24 items divided into 3 subscales: Pain (5 items), Stiffness (2 items), and Physical Function (17 items). Each item is on a 0-10 NRS scale. All questions are based on 48-hour recall. The pain questions are:

CCI [redacted]

The results of the questions on a 0-10 scale (0= no pain and 10=extreme pain) will be summed (0-50 scale for sum) and change from baseline (Day 1 of treatment period) to Day 4 will be measured across all treatment groups, and comparison of mean change from baseline will be tested using the ANCOVA model, with key pairwise comparisons.

8.3.7. Mean Change from Baseline to End of Treatment Period on the WOMAC Stiffness Subscale, 48-Hour Recall Version

The WOMAC stiffness subscale consists of 2 questions:

CCI [redacted]

The results of the questions on a 0-10 scale (0=no stiffness and 10=extreme stiffness) will be summed (0-20 scale for sum) and change from baseline (Day 1 of treatment period) to Day 4 will be measured across all treatment groups, and comparison of mean change from baseline will be tested using the ANCOVA model, with key pairwise comparisons.

8.3.8. Absolute WOMAC Stiffness Subscale Scores, 48-Hour Recall Version, at Each Time Point

The sum of WOMAC stiffness subscale results (0-20) will also be measured at each time point, and summarized descriptively by period and total in a table.

8.3.9. Mean Change from Baseline to End of Treatment Period on the WOMAC Physical Function Subscale, 48 Hour Recall Version

The WOMAC physical function subscale consists of 17 questions (see Appendix **Error! Reference source not found.** for details) and the results of the questions on a 0-10 scale (0=no difficulty and 10=extreme difficulty) will be summed (0-170 scale for sum) and change from baseline (Day 1 of treatment period) to Day 4 will be measured across all treatment groups, and comparison of mean change from baseline will be tested using the ANCOVA model, with key pairwise comparisons

8.3.10. Patient Global Impression of Change (PGIC) on Day 4 of Each Treatment Period

Approximately 8 hours after dosing, prior to leaving the study center on Day 4 of each Treatment Period, subjects will complete two Patient Global Impression of Change (PGIC) scales: the PGIC for OA Pain, and the PGIC for OA Symptoms. Each is a 7-point scale and the subject may only choose one response. The responses will be summarized in frequency tables and listed by subject.

8.3.11. Subject Preference for Treatment Assessed at the End of the Study

After treatment period 4 of 4 is completed, or if the subject is withdrawn from the study prior to completion, subjects will be asked to choose which treatment they felt gave the most pain relief. Responses will be summarized in frequency tables and listed by subject.

8.3.12. Level of Activities During Days 1 (Partial), 2 and 3 as Measured by the Accelerometer

The following activity level outcomes will be measured by the accelerometer:

- Number of steps per day
- Minutes active vs. inactive per day
- Percent time in low, medium or high activity
- Peak performance of the day (30 minutes)
- Maximum performance (1, 5, 20, 60 minutes)
- Cadence

Some endpoints mentioned in protocol (stride length, stride velocity, total distance) were not captured by the accelerometer and will not be presented. The results for each of these outcomes for the complete days (Day 2 and Day 3) will be summarized in a table and presented in a figure. Listings will be produced by subject.

9. Safety Analysis

Safety parameters include monitoring of TEAEs, physical examination findings, vital sign measurements, and laboratory examinations. All safety summaries and analyses will be summarized by actual treatment group within the Safety set.

9.1. Adverse Events

All AEs that occur during the study will be recorded and classified on the basis of Medical Dictionary for Regulatory Activities (MedDRA) terminology version 20.1. Treatment-emergent AEs (TEAEs) are defined as AEs with an onset on or after the date of the first IMP dose. Adverse events noted prior to the first IMP administration that worsen after baseline will also be reported as TEAEs and included in the summaries

All information pertaining to an AE noted during the study will be listed by subject, detailing verbatim term, preferred term, system organ class (SOC), onset date, resolution date, intensity, seriousness, action taken, outcome, drug relatedness, and relatedness to protocol-required procedures. The event onset will also be shown relative to IMP initiation (in number of days). Listings will be provided for all TEAEs.

TEAEs will be presented on the Safety set and displayed by treatment period and group unless noted.

9.1.1. Incidence of TEAEs

Adverse events will be deemed treatment emergent if the onset date is on or after the date of first treatment, Day 1 in each period. An overall summary of incidence of TEAEs (subject/event counts) including drug relation, serious adverse events (SAE), TEAEs leading to drug discontinuation and fatal AEs will be presented.

The number and percentage of subjects with at least one TEAE and the number of events overall will be presented by SOC and PT.

9.1.2. Relationship of Adverse Events to IMP

A summary of TEAEs by relationship to IMP will be presented in a table by total number of TEAE and incidence of occurrence. The investigator will provide an assessment of the relationship of the event to the IMP as Related or Not Related. If a subject reports multiple occurrences of the same TEAE, only the most closely related occurrence will be presented in the incidence count. Treatment-emergent AEs that are missing a relationship will be presented in the summary table as "Related" but will be presented in the data listing with a missing relationship. Percentages will be calculated based on the number of subjects in the safety set.

The TEAE data will be categorized and presented by SOC, PT, and relationship in a manner similar to that described in Section 9.1.1.

9.1.3. Intensity of Adverse Event

A summary of TEAEs by intensity will be presented in a table by total number of TEAE and incidence of occurrence. The intensity that will be presented represents the most extreme intensity captured on the eCRF page. The possible intensities are "Mild," "Moderate," and "Severe." In the TEAE intensity table, if a subject reported multiple occurrences of the same TEAE, only the most severe will be presented in the incidence count. Treatment-emergent AEs that are missing intensity will be presented in tables as "Severe" but will be presented in the data listing with a missing intensity. Percentages will be calculated based on the number of subjects in the safety set.

The TEAE data will be categorized and presented by SOC, PT, and intensity in a manner similar to that described in Section 9.1.1.

9.1.4. Serious Adverse Events

The treatment-emergent SAE data will be categorized and presented by SOC and PT in a manner similar to that described in Section 9.1.1. At each level of subject summarization, a subject is counted once for the incidence if the subject reported one or more events.

9.1.5. Deaths, Adverse Events Leading to Treatment Discontinuation/Study Discontinuation

A listing of all Serious TEAEs (including deaths) and/or AEs leading to treatment or study discontinuation will be presented. Additionally, a separate listing of any deaths will be presented. Deaths and discontinuations will be assessed based on data recorded in the AE eCRF page. Subject deaths will be identified as AEs where the outcome is "Fatal". Adverse events leading to treatment discontinuation will be identified as AEs where the action taken with IMP is "Drug Withdrawn". Adverse events leading to study discontinuation will be identified as AEs where the caused study discontinuation field is marked as "Yes".

9.2. Clinical Laboratory Evaluations

9.2.1. Hematology

The following hematology tests will be performed at screening: Hemoglobin, Hematocrit, RBC count, Platelet count, WBC count. The test results will be summarized and listed.

9.2.2. Serum Chemistry

The following serum chemistry tests will be performed at screening: BUN, Creatinine, Glucose, Ca++, Na+, K+, Cl-, Total CO2 (Bicarbonate), AST, ALT, Total Bilirubin, Alkaline phosphatase, Albumin, Total protein. The test results will be summarized and listed.

9.2.3. Urinalysis

The following urinalysis tests will be performed at screening: Specific gravity, pH, Glucose (qual), Protein (qual), Blood (qual), Ketones, Leukocyte Esterase, Nitrite, Bilirubin, Urobilinogen. The test results will be summarized and listed.

9.2.4. Urine Drug Screen

A screen for drugs of abuse will be performed at Screening (Visit 1) and Visit 2 of each treatment period. A by-subject list of urine drug screen will be presented. Subjects will be tested for the following drugs: methamphetamines, amphetamines, cannabinoids, cocaine, opiates, benzodiazepines, barbiturates, methylenedioxymethamphetamine, methadone, oxycodone, adulterants.

9.2.5. Breath/Saliva Alcohol Test

An alcohol test, administered via breath or saliva will be performed at Screening and Visit 2 of each treatment period for inclusion/exclusion purposes. A by-subject listing of results will be presented.

9.2.6. Urine Pregnancy Test

At screening and Visit 2 of each treatment period all females of child bearing potential must perform a urine pregnancy test and verify that the results are negative (not pregnant). Subjects with a verified positive pregnancy test will be withdrawn from the study. Pregnancy test results will be listed by subject.

9.3. Vital Sign Measurements

The vital sign summary and analysis will be based on the recordings of the variables blood pressure (systolic and diastolic [mm Hg]), oral body temperature (F), heart rate (beats/min), and respiratory rate (breaths/min). Notable abnormal values will be flagged in the data listings. The conversion between conventional unit to SI unit for temperature and weight and definitions for notable abnormal values are listed in the appendix (Section **Error! Reference source not found.**). After screening assessment, vital signs assessments will consist of heart rate and blood pressure only. Observed values at Screening and Visits 2 and 3 and change from Baseline values to Visit 3 for vital sign measurements will be summarized descriptively by treatment group.

All summaries will be done for the Safety set and all vital sign data will be listed.

9.4. Physical Examination

The physical examination (by means of inspection, palpation, auscultation) will be performed by a physician (or appropriately licensed designee) at the study site covering at least the organs of the cardiovascular, respiratory, and abdominal system.

Abnormal physical examination findings are recorded either as medical history or as adverse events. Physical examination results will be listed.

10. Interim Analysis

There is no interim analysis scheduled for this study.

11. References

1. Altman, R. Osteoarthritis (OA). In Merck manual online. Retrieved from http://www.merckmanuals.com/home/bone,-joint,-and-muscle-disorders/joint-disorders/osteoarthritis-%28oa%29 Accessed on April 11, 2016.

2. Centers for Disease Control and Prevention (CDC). Prevalence of doctor diagnosed arthritis and arthritis attributable activity limitation—United States, 20072009. *Morb Mortal Wkly Rep*. 2013; 59(39):1261-5.

3. Zhang Y, Jordan, M. Epidemiology of Osteoarthritis. *Clin Geriatr Med.* 2010; 26(3): 355-369.

4. Bannuru, R. et al. Comparative Effectiveness of Pharmacologic Interventions for Knee Osteoarthritis. *Ann Intern Med*, 2015; 162:46-54.

5. Altman, R. et al. Three-Month Efficacy and Safety of Acetaminophen Extended-Release for Osteoarthritis Pain of the Hip or Knee: A Randomized, Double-Blind, Placebo-Controlled Study. *Osteoarthritis Cartilage*. 2007; 15, 454-461.

6. Golden, HE, Moskowitz, RW, and Minic, M. Analgesic efficacy and safety of nonprescription doses of naproxen sodium compared with acetaminophen in the treatment of osteoarthritis of the knee. *Am J Ther*. 2004;11:85-94.

7. Schiff, M and Minic, M. Comparison of the analgesic efficacy and safety of nonprescription doses of naproxen sodium and Ibuprofen in the treatment of osteoarthritis of the knee. *Journal of Rheumatology*. 2004;31:1373-1383.

8. Cuoto A, Troullos E, Moon J, Paredes-Diaz A. Analgesic efficacy & safety of nonprescription doses of naproxen sodium in the treatment of osteoarthritis of the knee or hip. *Osteoarthritis and Cartilage* 2017;25 (supplement 1): S420 Abstract #692.

9. Puljak L, Marin A, Vrdoljak D, Markotic F, Utrobicic A, Tugwell P. Celecoxib for osteoarthritis. *Cochrane Database of Systematic Reviews* 2017, Issue 5. Art. No.: CD009865. DOI: 10.1002/14651858.CD009865.pub2.

10. Bellamy N, Buchanan W, Goldsmith C, Campbell J, Stitt L. Validation Study of WOMAC: A Health Status Instrument for Measuring Clinically Important Patient Relevant Outcomes to Antirheumatic Drug Therapy in Patients with Osteoarthritis of the Hip or Knee. J Rheumatol. 1988 Dec; 15(12):1833-40.

11. Dixon SJ, Hinman RS, Creaby MW, Kemp G, Crossley KM. Knee Joint Stiffness During Walking in Knee Osteoarthritis. *Arthritis Care Res*. 2010 Jan 15;62(1):38-44.

12. Ghomrawi HM, Mandl LA, Rutledge J, Alexiades MM, Mazumdar M. Is there a role for expectation maximization imputation in addressing missing data in research using WOMAC questionnaire? Comparison to the standard mean approach and tutorial. *BMC Musculoskeletal Disorders*. 2011; 12(109): DOI: 10.1186/1471-2474-12-109.

13. Kenward MG, Roger JH. The use of baseline covariates in crossover studies. *Biostatistics*, Jan 2010; 1-17.

14. Choi K, Hong T, Lee J. On comparison of SAS Codes with GLM and MIXED for the crossover studies with QT interval data. Transl Clin Pharamcol, 2014; 78-82.

15. Yarandi, HN. Paper SD04 – Crossover Designs and Proc Mixed in SAS. http://analytics.ncsu.edu/sesug/2004/SD04-Yarandi.pdf

12. Appendices

12.1. Notable Abnormal Vital Signs

Table 2 Notable Abnormal Vital Signs

Parameter	Actual Value	Change from Baseline
Pulse	< 40 bpm > 120 bpm	>= 15 bpm increase from baseline >= 15 bpm decrease from baseline
Systolic blood pressure	< 80 mmHg > 180 mmHg	>= 20 mmHg increase from baseline >= 20 mmHg decrease from baseline
Diastolic blood pressure	< 50 mmHg > 105 mmHg	>= 15 mmHg increase from baseline >= 15 mmHg decrease from baseline
Temperature	NA	>= 2 F increase from baseline >= 2 F decrease from baseline

12.2. Schedule of Study Procedures

Trial Procedure	Visit 1 Screening	Screening Washout Phase	Treatment Phase These events are repeated during each of the 4 Treatment Periods. Treatment Periods will be separated by a washout of 3 to 7 days						End of Study (Phone call)
			Visit 2 Start of Treatment Period		Continue Treatment Period Outsubject	Visit 3 End of Treatment Period			
	Day -18 to -5	Day -4 to -1 (+3 d)	Day 1	2, 4, 6, 8 hrs	Day 2-3	Day 4	2, 4, 6 hrs	8 hrs	7-10 days after Treatment Period 4 of 4
Signed Informed Consent	X								
Inclusion/Exclusion criteria review	X		X			X			
Subject demographics	X								
Medical history	X								
Prior/Concomitant medications	X		X			X			X
History of drug, alcohol and tobacco use	X								
Body weight, height and BMI	X								
Physical examination	X								

OA assessment of the knee (includes radiograph)	X								
Vital signs[a]	X		X			X			
Clinical Laboratory Tests (hematology, chemistry, urinalysis)	X								
Alcohol breath/saliva test	X		X						
Urine drug screen	X		X						
Urine pregnancy test (if applicable)	X		X						
Accurate pain reporting training	X		X						
Appropriate expectations training	X		X						
Regional Pain Scale	X								
48 hour WOMAC Pain	X		X[b]			X			
48 hour WOMAC Stiffness	X		X			X			
48 hour WOMAC Function	X		X			X			
24 hour Average Pain Intensity NRS	X		X			X			
Baseline Stiffness questionnaire	X								
Joint Stiffness Severity NRS	X								
Washout from analgesic medications		X						X[d]	
Daily BASS			X		X	X			
Accelerometer Setup				X					
Accelerometer Use				X	X				
Accelerometer Data Download						X			
Accelerometer Battery Charge						X			
Current Pain Intensity NRS			X[c]			X[c]			
STEPP →10 point pain NRS	X[c]		X[f]	X[f]		X[f]	X[f]	X[f]	
Randomization			X[g]						
IMP self-administration			X		X	X[h]			
IMP compliance review			X			X			
Subject Global Impression of Change								X	
Subject treatment preference								X[i]	
Qualitative interview								X[i]	
Adverse events			X	X	X	X	X	X	X

[a] Vital signs (blood pressure, respiratory rate, heart rate and body temperature after 5 minutes of rest in a sitting position). After the screening assessment, vital signs assessments will consist of heart rate and blood pressure

[b] WOMAC pain subscale average score must be ≥4.0 on 0-10 scale to be eligible for Randomization on Day 1 of Treatment Period 1 of 4

[c] Screening STEPP is to confirm subject is capable of performing the exercise and completing the pre-STEPP and post-STEPP NRS

[d] A washout period of 3 to 7 days will occur at the start of Treatment Periods: 2 of 4, 3 of 4, and 4 of 4. During each washout, only acetaminophen ER will be allowed as rescue therapy (as needed up to 2000 mg/day)

[e] Current pain intensity performed immediately prior to the pre-dose STEPP and hourly at 1, 3, 5 and 7 hours after dosing

[f] Subjects will complete the 0-10-point NRS pain scale pre-dose (time 0) and immediately before (at rest) and then 2-5 minutes after completing the STEPP at 2, 4, 6 and 8 hours post-dose. Additionally, subjects will complete hourly current pain intensity rating 1, 3, 5, and 7 hours post-dose when not completing the STEPP

[g] Subjects will be randomized to a treatment sequence at Treatment Period 1 of 4

[h] Subjects will only take the 8 AM dose that day

[i] Completed at Visit 3, Treatment Period 4 of 4 only or if the subject is withdrawn prior to completion

12.3. WOMAC Function Subscale Questions

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]