STATISTICAL ANALYSIS PLAN SM04690-OA-04
(Version 01, 13 September 2018)

A Phase 2, 24-Week, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Safety and Efficacy of SM04690 for the Treatment of Moderately to Severely Symptomatic Knee Osteoarthritis

NCT03122860

STATISTICAL ANALYSIS PLAN SM04690-OA-04

Study Title:	A Phase 2, 24-Week, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Safety and Efficacy of SM04690 for the Treatment of Moderately to Severely Symptomatic Knee Osteoarthritis
Protocol Number:	SM04690-OA-04
Protocol Version:	Amendment 03 Version 00, April 23, 2018
SAP Version:	01
Investigational Product:	SM04690 Injectable Suspension
Clinical Phase:	Phase 2

Sponsor: Samumed, LLC.
9381 Judicial Dr., Suite 160
San Diego, CA 92121
(858) 926-2900

SAP SIGNATURE PAGE

A Phase 2, 24-Week, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Safety and Efficacy of SM04690 for the Treatment of Moderately to Severely Symptomatic Knee Osteoarthritis

Version: 01

Date: September 13, 2018

Name & Title	Signature	Date
[redacted]		13 Sep 2018
[redacted]		13 SEP 2018
[redacted]		13 SEP 2018

Samumed commits to satisfying the requirements of the ICH-GCP Guidelines regarding the responsibilities of the Sponsor, the US Code of Federal Regulations 21 CFR parts 50, 54, 56, 312, and 314 and Good Clinical Practice Guidelines, as applicable.

TABLE OF CONTENTS

SAP SIGNATURE PAGE 2
TABLE OF CONTENTS 3
1. INTRODUCTION 5
2. STUDY OBJECTIVES 5
2.1. Primary Objective 5
2.2. Secondary Objective 5
3. STUDY DESIGN 5
3.1. Sample Size 5
3.2. Randomization 5
3.3. Study Medication Dosing 6
4. STUDY ENDPOINTS 6
4.1. Primary Endpoints 6
4.2. Secondary Endpoint 6
4.3. Exploratory Endpoints 7
5. STUDY OUTCOME MEASURES 7
5.1. Safety Outcome Measures 7
5.2. Efficacy Outcome Measures 7
5.2.1. Pain Numeric Rating Scale 8
5.2.2. Western Ontario and McMaster Universities Arthritis Index 8
5.2.3. Knee Injury and Osteoarthritis Outcome Score 8
5.2.4. Knee Injury and Osteoarthritis Outcome Score – Short Form 8
5.2.5. Patient Global Assessment of Disease Activity 9
5.2.6. Physician Global Assessment of Disease Activity 9
5.2.7. Nonsteroidal Anti-Inflammatory Drug Usage 9
5.2.8. Evaluation of Success of Blinding 9
5.2.9. Joint Space Width 9
6. ANALYSIS DATASETS 9
6.1. Full Analysis Set 9
6.2. Modified Full Analysis Set 10
6.3. Per-Protocol Analysis Set 10
6.4. Safety Analysis Set 10

7. EFFICACY ANALYSIS 10
7.1. General Considerations 10
7.2. ANCOVA Model for Continuous Efficacy Outcomes 10
7.3. Analysis of Primary Efficacy Endpoints 11
7.4. Analysis of Secondary Efficacy Endpoints 11
7.5. Analysis of Exploratory Efficacy Endpoints 11
7.5.1. Additional Timepoints and Outcome Measures 11
7.5.2. Mixed-Effects Model for Repeated Measures 11
7.5.3. Subject Perception of Treatment 12
7.5.4. Comparison of Vehicle and Sham 12
7.5.5. Subgroup Analysis 12
7.5.6. Concordance Analysis 13
7.6. Dose Selection 13
8. SAFETY ANALYSIS 13
8.1. Adverse Events 13
8.2. Clinical Laboratory 13
8.3. Vital Signs 14
8.4. Concomitant Medications 14
8.5. Extent of Exposure 14
9. STATISTICAL CONSIDERATIONS 14
9.1. Baseline 14
9.2. Early Termination 14
9.3. Electronic Device Entries 14
9.4. Handling of Missing Data 15
10. STUDY SUBJECTS 15
10.1. Disposition of Subjects 15
10.2. Demographic and Baseline Characteristics 15
10.3. Medical History 15
11. PROTOCOL DEVIATIONS 16
12. ANALYSIS SOFTWARE 16
13. REFERENCES 17

1. INTRODUCTION

Osteoarthritis (OA) is the most common form of arthritis and causes degenerative structural change that results in pain and decreased mobility. Samumed, LLC is developing SM04690, a small-molecule inhibitor of the Wnt pathway, for the treatment of OA. The purpose of this study is to assess the safety and efficacy of four different strengths of SM04690 administered by intra-articular injection into the target knee joint of moderately to severely symptomatic OA subjects.

2. STUDY OBJECTIVES

2.1. Primary Objective

The primary objective of this study is to determine the effective dose(s) of SM04690 for the treatment of knee OA.

2.2. Secondary Objective

The secondary objective of this study is to evaluate the safety and tolerability of SM04690.

3. STUDY DESIGN

This study is a multicenter, randomized, double-blind, placebo-controlled, parallel-group study of four different concentrations of SM04690 injected into the target knee joint of moderately to severely symptomatic OA subjects.

3.1. Sample Size

A sample size of approximately 630 subjects was randomized at a ratio of 1:1:1:1:1:1 (0.03 mg active per 2 mL injection : 0.07 mg active per 2 mL injection : 0.15 mg active per 2 mL injection : 0.23 mg active per 2 mL injection : 0.0 mL vehicle : 2.0 mL vehicle). The sample size for this study was based upon accepted dose-finding statistical practice (Ting et al. 2017).

3.2. Randomization

Subjects were randomized to each treatment group using a permuted block design. Prior to randomization, the presence of unilateral or bilateral knee OA was determined by the investigator and subjects were required to complete the Widespread Pain Index and Symptom Severity (WPI&SS) questionnaire.

The Widespread Pain Index (WPI) is a body map consisting of 19 prespecified areas where subjects can indicate the presence of pain during the past seven days with a possible score 0-19. Question 2 of the Symptom Severity section of the questionnaire (SSQ2) focuses on the presence and severity of three prespecified symptoms (1. fatigue, 2. trouble thinking or remembering, 3. waking up tired) during the past seven days with a possible score of 0-9.

For permuted block randomization, subjects were stratified by knee OA laterality (50% unilateral, 50% bilateral) and WPI&SS score (80% WPI ≤ 4 and SSQ2 ≤ 2, 20% WPI > 4 and/or SSQ2 > 2). A randomization list with a block size of 6 was generated and stored in Medidata Balance. The block size of 6 was selected to accommodate the 6 possible treatment assignments

and limit the possibility of imbalance. Since the study was not stratified by site, predictability with a small block size was not a concern due to competitive enrollment across all sites.

Details on the configuration of Medidata Balance for randomization design and list generation were documented in the Randomization Configuration Plan. The seed used to generate the randomization list was randomly assigned and captured in the Medidata Balance audit trail. The randomization lists and seeds were not accessible to blinded study personnel.

3.3. Study Medication Dosing

SM04690 will be administered in the following dosage strengths:

- SM04690 0.03 mg in 2-mL Injectable Suspension
- SM04690 0.07 mg in 2-mL Injectable Suspension
- SM04690 0.15 mg in 2-mL Injectable Suspension
- SM04690 0.23 mg in 2-mL Injectable Suspension
- SM04690 0 mg; 2 mL vehicle injection only
- SM04690 0 mg; 0 mL sham injection (intra-articular injection procedure is performed identically to all other treatment groups but without the injection of any liquid)

4. STUDY ENDPOINTS

4.1. Primary Endpoints

The primary endpoints of this study include:

1. Evaluate change from baseline OA pain in the target knee as assessed by the weekly averages of daily pain Numeric Rating Scale (Pain NRS) at Week 24
2. Evaluate change from baseline OA pain in the target knee as assessed by Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain subscore at Week 24
3. Evaluate change from baseline OA function in the target knee as assessed by WOMAC Physical Function (WOMAC Function) subscore at Week 24
4. Evaluate change from baseline in medial joint space width (mJSW) as documented by radiograph of the target knee at Week 24

4.2. Secondary Endpoint

The secondary endpoints of this study include:

1. Evaluate the safety of SM04690 as measured by treatment-emergent adverse events (TEAEs)
2. Evaluate change from baseline OA disease activity as assessed by Patient Global Assessment (PTGA) at Week 24

4.3. Exploratory Endpoints

Additional exploratory endpoints of this study include:

1. Evaluate change from baseline OA pain in the target knee as assessed by WOMAC Pain at Weeks 4, 8, 12, 16, and 20
2. Evaluate change from baseline OA function in the target knee as assessed by WOMAC Function at Weeks 4, 8, 12, 16, and 20
3. Evaluate change from baseline symptoms of OA in the target knee as assessed by WOMAC Total at Weeks 4, 8, 12, 16, 20, and 24
4. Evaluate change from baseline OA pain in the target knee at each week (aside from Week 24) as assessed by Pain NRS
5. Evaluate change from baseline OA disease activity as assessed by PTGA at Weeks 4, 8, 12, 16, and 20
6. Evaluate change from baseline OA function in the target knee as assessed by Knee Injury and Osteoarthritis Outcome Score (KOOS) at Weeks 4, 8, 12, 16, 20, and 24
7. Evaluate change from baseline OA function in the target knee at each week as assessed by KOOS Physical Function Short Form (KOOS-PS)
8. Evaluate change from baseline OA disease activity as assessed by Physician Global Assessment (PGA) at Weeks 4, 12, and 24
9. Evaluate change over time in NSAID usage as assessed by electronic diary responses at every week
10. Evaluate subject perception of study treatment received at Day 1 and Week 24
11. Evaluate outcome differences between treatment with sham and vehicle injections at Weeks 4, 8, 12, 16, 20 and 24

5. STUDY OUTCOME MEASURES

5.1. Safety Outcome Measures

Safety will be assessed by summarizing and evaluating TEAEs, clinical laboratory results, vital signs and concomitant medications.

5.2. Efficacy Outcome Measures

Knee pain was assessed by the subject using Pain NRS. Knee pain and function was assessed by the subject using the WOMAC, KOOS and KOOS-PS questionnaires. Disease activity was assessed by the PTGA and PGA. NSAID usage was assessed by subject self-report of taking an NSAID via electronic diary. The subjects' perception of study treatment was assessed by a paper questionnaire administered immediately after treatment and again at the end of the study. JSW was assessed with radiograph imaging of the target knee.

5.2.1. Pain Numeric Rating Scale

The Pain NRS is an 11-point scale (0-10) for subject self-reporting of average knee pain in the last 24 hours, where 0 represents no pain and 10 represents extreme pain. Subjects were prompted to report average pain in the target knee daily on their electronic device (between 5:00 pm and 11:59 pm) from Screening Visit 2 through the end of the study. An average weekly score (referred to as weekly Pain NRS) was calculated for each subject if they had provided a response for at least 4 out of 7 days in a given week. Weeks are defined as Day -7 through Day -1 (day before treatment) for baseline, Day 1 (day of treatment) through Day 7 for Week 1, Day 8 through Day 14 for Week 2, and so on.

5.2.2. Western Ontario and McMaster Universities Arthritis Index

The WOMAC Version NRS 3.1 instrument is a patient-reported outcome measure to assess the symptoms of OA. Subjects were prompted to complete the WOMAC instrument monthly on their electronic device. The devices required subjects to enter a response for each item before proceeding to the next.

WOMAC consists of 24 questions in three domains: physical function (17 questions), pain (5 questions) and stiffness (2 questions). The response for each question in the NRS format ranges from 0 to 10. Each domain subscore as well as a total score are calculated by adding together the numerical responses for a range of 0 to 240 total points. For analysis, WOMAC Total score as well as all three subscores will be linearly transformed to a 0-100 scale, where 0 represents no difficulty and 100 represents extreme difficulty.

5.2.3. Knee Injury and Osteoarthritis Outcome Score

KOOS is a widely used, non-proprietary tool intended to evaluate pain, symptoms, 2 types of physical function and knee-related quality-of-life. Subjects were prompted to complete the KOOS questionnaire monthly on their electronic device. The devices required subjects to enter a response for each item before proceeding to the next.

Subjects were instructed to answer each item with one of 5 standardized Likert responses. Each item is assigned a score of 0-4 and a normalized score of 0-100 is calculated for each of the 5 subscales (Pain, Symptoms, Activities of Daily Living, Sport and Recreation, Quality of Life) based on the August 2012 scoring instructions (KOOSscoring2012.pdf). Per these instructions, two composite scores are also calculated: $KOOS_5$ (the average of all 5 subscales) and $KOOS_4$ (the average of 4 subscales, excluding ADL). For all subscale and composite scores, 100 represents no difficulty and 0 represents extreme difficulty.

5.2.4. Knee Injury and Osteoarthritis Outcome Score – Short Form

KOOS-PS is a 7-item measure that is used to assess physical function. Subjects were prompted to complete the KOOS-PS questionnaire weekly on their electronic device. The devices required subjects to enter a response for each item before proceeding to the next.

Subjects were instructed to answer each item with one of 5 standardized Likert responses. Each item is assigned a score of 0-4 and a normalized score of 0-100 is calculated. For this study, to align with KOOS, 100 represents no difficulty and 0 represents extreme difficulty.

5.2.5. Patient Global Assessment of Disease Activity

The patient assessment was a 50 mm visual analog scale (VAS) adapted from the Patient Assessment Form © 1999, American College of Rheumatology. Subjects rated how well they were doing, considering all the ways in which illness and health conditions may affect them. The VAS was anchored by "Very Well" (0) on the left and "Very Poorly" (50) on the right. Subjects were prompted to complete the assessment monthly on their electronic device. Scores were subsequently scaled to 0-100.

5.2.6. Physician Global Assessment of Disease Activity

The physician assessment was a 100 mm VAS adapted from the Physician Assessment Form © 1999, American College of Rheumatology. The investigator rated the subject's disease activity on a paper form on Day 1 and Weeks 4, 12 and 24. The VAS was anchored by "Very Good" (0) on the left and "Very Bad" (100) on the right.

5.2.7. Nonsteroidal Anti-Inflammatory Drug Usage

Subjects were asked "Did you take any medication that is an NSAID in the last 24 hours?" on their electronic device daily (between 5:00 pm and 11:59 pm) from Screening Visit 2 through the end of the study. If a "yes" or "no" response was provided for at least 4 out of 7 days in a given week, a proportion of days with NSAID usage is calculated (days with "yes" response / total number of responses). Weeks are defined as Day -7 through Day -1 for baseline, Day 1 through Day 7 for Week 1, Day 8 through Day 14 for Week 2, and so on.

5.2.8. Evaluation of Success of Blinding

Subjects were prompted, "Please indicate which treatment arm you believe you were assigned to for this study" on a paper form on Day 1 (post-injection) and on Week 24. The form captured responses of "Study drug injection," "Injection of 2 mL inactive vehicle substance," "Needle insertion into the knee with no vehicle substance injected" and "Do not know".

5.2.9. Joint Space Width

Radiograph of the knee joints was taken at Screening Visit 1 and Week 24. Whenever possible, the radiographs obtained the posterior-anterior (PA) view according to the Image Review Charter. Medial and lateral JSW measurements of the target knee were provided by the vendor, Medical Metrics, Inc. (MMI).

6. ANALYSIS DATASETS

6.1. Full Analysis Set

The Full Analysis Set (FAS) includes all subjects who were randomized and received a study injection.

6.2. Modified Full Analysis Set

The Modified Full Analysis Set (mFAS) includes FAS subjects who received a protocol-specified injection. Subjects who were administered a treatment that is not prescribed by the protocol are excluded from this analysis set.

6.3. Per-Protocol Analysis Set

The Per-Protocol Analysis Set (PPAS) includes mFAS subjects who completed the study and did not have any major protocol deviations (see Section 11).

6.4. Safety Analysis Set

The Safety Analysis Set (SAS) includes all subjects who received a study injection.

7. EFFICACY ANALYSIS

7.1. General Considerations

Unless otherwise specified, efficacy analyses will be performed on the FAS, mFAS and PPAS. For continuous variables within each treatment group, the outcome measure at each visit, as well as absolute change (outcome – baseline), will be summarized using descriptive statistics (number of subjects, mean, standard deviation, median, minimum and maximum). Categorical variables will be summarized with frequency tables.

Since this study was stratified based upon known clinical phenotypes of OA, all efficacy analyses will be completed in three subject populations: 1) subjects with unilateral symptomatic OA, 2) subjects with unilateral symptomatic OA and without widespread pain (WPI ≤ 4 and SSQ2 ≤ 2), and 3) all subjects.

Subject-level listings will be provided for each outcome measure.

7.2. ANCOVA Model for Continuous Efficacy Outcomes

ANCOVA models will be used for all continuous efficacy outcome measures summarized by change from baseline in order to test the following four hypotheses:

$$H_0: (\beta_i - \beta_0) = 0$$

$$H_A: (\beta_i - \beta_0) \neq 0, \quad \text{where } i = 1,2,3,4$$

In the statement above, β is the least squares estimate in the change in the continuous efficacy outcome from baseline at each timepoint (where β_0 is the estimate for vehicle), and i represents each of the four SM04690 treatment groups.

Least squares estimate of difference between each treatment group and placebo in the change in the continuous efficacy outcome from baseline at each timepoint, adjusted for baseline value, will be reported along with unadjusted 95% confidence intervals and P values.

7.3. Analysis of Primary Efficacy Endpoints

The change from baseline in weekly Pain NRS, WOMAC Pain, WOMAC Function and mJSW at Week 24 will be analyzed with the ANCOVA model described in Section 7.2. Figures displaying the change in outcome over time will be provided for the FAS and mFAS.

7.4. Analysis of Secondary Efficacy Endpoints

The change from baseline in PTGA at Week 24 will be analyzed with the ANCOVA model described in Section 7.2. Figures displaying the change in PTGA over time will be provided for the FAS and mFAS.

7.5. Analysis of Exploratory Efficacy Endpoints

7.5.1. Additional Timepoints and Outcome Measures

The change from baseline in the parameters described below will be analyzed with the ANCOVA model described in Section 7.2.

1. WOMAC Pain at Weeks 4, 8, 12, 16, and 20
2. WOMAC Function at Weeks 4, 8, 12, 16, and 20
3. WOMAC Total at Weeks 4, 8, 12, 16, 20, and 24
4. Weekly Average of Pain NRS at all weeks, except Week 24
5. PTGA at Weeks 4, 8, 12, 16, and 20
6. KOOS at Weeks 4, 8, 12, 16, 20, and 24
7. KOOS-PS at all weeks
8. PGA at Weeks 4, 12, and 24

7.5.2. Mixed-Effects Model for Repeated Measures

Change over time in primary and secondary endpoints (weekly Pain NRS, WOMAC Pain, WOMAC Function and PTGA) as well as weekly measures (KOOS-PS and NSAID usage) will be characterized using mixed-effects models for repeated measures (MMRM). The models will estimate change from baseline with treatment, week, treatment×week interaction and baseline value as covariates.

For weekly measured outcomes, MMRM will be estimated assuming a Toeplitz variance-covariance matrix. The Toeplitz structure is a more generalized form of autogressive-1 (AR1) structure, allowing the data to inform how the correlation between within-subject observations decreases over time instead of implicitly defining a uniform decay structure. For monthly measured outcomes, MMRM will be estimated assuming an unstructured variance-covariance matrix, allowing the data to fully inform the correlation between within-subject observations. (Kincaid 2005)

7.5.3. Subject Perception of Treatment

The following contingency table will be prepared to evaluate the subjects' perception of treatment received at Day 1 and at Week 24.

Assignment	Guess				Total
	SM04690 (1)	Vehicle (2)	Sham (3)	Don't Know (4)	
SM04690 (1)	n_{11} ($P_{1\|1}$)	n_{12} ($P_{2\|1}$)	n_{13} ($P_{3\|1}$)	n_{14} ($P_{4\|1}$)	$n_{1.}$
Vehicle (2)	n_{21} ($P_{1\|2}$)	n_{22} ($P_{2\|2}$)	n_{23} ($P_{3\|2}$)	n_{24} ($P_{4\|2}$)	$n_{2.}$
Sham (3)	n_{31} ($P_{1\|3}$)	n_{32} ($P_{2\|3}$)	n_{33} ($P_{3\|3}$)	n_{34} ($P_{4\|3}$)	$n_{3.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	N

$P_{j|i}$ = P(guess j|assigned treatment i) = probability of guessing treatment j when assigned treatment i

Bang's Blinding Index (BI) (Bang et al. 2010) will be calculated to determine the percentage of unblinding that is beyond chance.

$$\text{Bang's BI}_i = \left(2\frac{n_{ii}}{n_{i1}+n_{i2}+n_{i3}} - 1\right) \times \left(\frac{n_{i1}+n_{i2}+n_{i3}}{n_{i1}+n_{i2}+n_{i3}+n_{i4}}\right),$$

where $BI_i = 0$ represents random guessing; $BI_i = 1$ represents complete unblinding; $BI_i = -1$ represents opposite guessing

7.5.4. Comparison of Vehicle and Sham

The change from baseline in weekly Pain NRS, WOMAC Pain, WOMAC Function and PTGA at Weeks 4, 8, 12, 16, 20 and 24, as well as mJSW at Week 24 will be analyzed with an ANCOVA model with the following hypothesis:

$$H_0: (\beta_{vehicle} - \beta_{sham}) = 0 \; ; \; H_A: (\beta_{vehicle} - \beta_{sham}) \neq 0$$

Least squares estimate of difference between vehicle and sham in the change in outcome from baseline, adjusted for baseline value, will be reported along with unadjusted 95% confidence intervals and P values.

7.5.5. Subgroup Analysis

Efficacy analyses described in Sections 7.3 and 7.4 will be further analyzed by additional, complementary clinical phenotypes of OA not already defined in Section 7.1.

The additional clinical phenotypes of OA include:

1. Bilateral Symptomatic OA
2. WPI ≤ 4 and SSQ2 ≤ 2
3. WPI > 4 and/or SSQ2 > 2
4. Unilateral Symptomatic OA with WPI > 4 and/or SSQ2 > 2
5. Bilateral Symptomatic OA with WPI ≤ 4 and SSQ2 ≤ 2

6. Bilateral Symptomatic OA with WPI > 4 and/or SSQ2 > 2

7.5.6. Concordance Analysis

Logistic regression will be conducted within each treatment group to determine whether baseline-adjusted change in mJSW is concordant with response criteria for Pain NRS, WOMAC Pain, WOMAC Function, and Patient Global. Response criteria will be patterned after the OMERACT-OARSI response criteria (Pham 2003); a responder must have a 50% improvement with corresponding 20 point (out of 100) actual scale improvement. Each outcome will be modeled individually, and OMERACT-OARSI 'strict' responder (response on either WOMAC Pain or WOMAC Function) as well as Pain and Function responder (response on both WOMAC Pain and WOMAC Function) will also be modeled. Receiver operator characteristic (ROC) curves and ROC areas under the curve (AUC) will be provided.

7.6. Dose Selection

Multiple Comparison Procedure Modelling (MCP-Mod) will be used to estimate SM04690 dose response compared to vehicle using the Week 24 baseline-adjusted ANCOVA analysis of the primary and secondary outcome measures in the FAS and all three OA phenotypes described in Section 7.1. MCP-Mod will use both monotonic (e.g. linear) and non-monotonic (e.g. beta-mod, quadratic, emax) models in exploring outcomes' dose response relationship.

8. SAFETY ANALYSIS

The analysis of safety outcome measures will be performed on the SAS.

8.1. Adverse Events

All adverse events (AEs) collected in this study are TEAEs. AEs will be presented in summary tables depicting the number of AEs and the number and percent of unique subjects experiencing each AE within each treatment group. The following summaries will be provided:

- AEs by seriousness, toxicity grade severity and relationship to study product
- AEs by MedDRA system organ class and preferred term
- AEs by preferred term, sorted by prevalence

Separate subject-level listings will be provided for all serious and non-serious AEs.

8.2. Clinical Laboratory

All chemistry, hematology and urinalysis results from the central lab will be summarized into shift tables as normal, non-clinically significant abnormal, and clinically significant abnormal. Assessments of clinical significance for abnormal values were made by the investigator on results that were outside of the normal range or had a toxicity grade of 1 or greater. Shift tables will compare the number and percent of assessments from each visit to baseline values for each treatment group. Abnormal results for each subject will be provided in listings that will include assay name, result, normal range and an explanation for clinically significant values.

Abnormal urine microscopy and manual differential assessments will be listed for each subject.

8.3. Vital Signs

Weight, height, body mass index (BMI) and vital signs, including systolic blood pressure, diastolic blood pressure, pulse rate, respiratory rate and body temperature, will be summarized for each treatment group. A statistical description (number of subjects, mean, standard deviation, median, minimum and maximum) of each parameter at baseline will be provided along with the change from baseline at each subsequent visit. A subject-level listing will also be provided.

8.4. Concomitant Medications

The World Health Organization Drug Dictionary (WHODD) will be used to classify prior and concomitant medications by Anatomical Main Group (Anatomical Therapeutic Chemical, ATC, Level 1), Therapeutic Subgroup (ATC Level 2), and preferred term. Prior and concomitant medication usage will be summarized by the number and percentage of subjects receiving each medication by treatment group.

Subject-level listings containing prior and concomitant medications (WHODD coding), and procedures and non-drug therapies (MedDRA coding) will be provided.

8.5. Extent of Exposure

All treated subjects received a single injection. A list of lot numbers used in this study will be provided in the clinical study report.

9. STATISTICAL CONSIDERATIONS

9.1. Baseline

Baseline is defined as the last value recorded for any given parameter prior to study medication injection. If a subject never received a study injection, baseline is defined as the last value recorded prior to study termination.

9.2. Early Termination

If a subject discontinues the study, early termination assessments will be performed according to the protocol. If these assessments occur within the window of a scheduled visit (+3 days for Weeks 4 and 12 and +7 days for Week 24), they will be associated with that visit for the purposes of FAS and mFAS analysis.

9.3. Electronic Device Entries

Section 7.3.7 of the Protocol provides the schedule of electronic diary and questionnaire completion. Weeks are defined as Day -7 through Day -1 (day before treatment) for baseline, Day 1 (day of treatment) through Day 7 for Week 1, Day 8 through Day 14 for Week 2, and so on. If there is more than one response to any parameter during the associated time period (e.g. more than one Pain NRS entry in one day or more than one KOOS-PS entry in one week), the following steps will be taken in the presented order:

1. Weekly or monthly questionnaires that were completed on a day not specified in Section 7.3.7 of the Protocol (due to device error) will not be considered for analysis
2. Any response after the initial response will not be considered for analysis

9.4. Handling of Missing Data

Missing data will not be imputed for the efficacy analyses described in Section 7. A sensitivity analysis will be performed using the last observation carried forward (LOCF) approach to impute missing assessments at Week 24 for weekly Pain NRS, WOMAC Pain, WOMAC Function and PTGA. This will then be followed by the ANCOVA analysis described in Section 7.2. Since single imputation is not necessary for the PPAS with complete data, LOCF analysis will be performed on the FAS and mFAS, and repeated for the OA phenotypes described in Section 7.1.

In addition, a second sensitivity analysis for change from baseline in weekly Pain NRS, WOMAC Pain, WOMAC Function and PTGA will be conducted using MMRM as outlined in Section 7.5.2.

10. STUDY SUBJECTS

10.1. Disposition of Subjects

Subject disposition will be presented in a summary table detailing the number and percentage of subjects who were consented, randomized, treated, completed the study or discontinued (e.g. screen failure, subject decision, etc.) by treatment group, site and randomization cohort. The disposition for individual subjects will be listed along with additional information on discontinued subjects.

10.2. Demographic and Baseline Characteristics

Demographic and baseline characteristics, including gender, race, ethnicity, age, weight, height, body mass index (BMI), Kellgren-Lawrence (KL) grade, OA laterality, WPI and SSQ2 will be presented by treatment group. Continuous variables will be summarized with descriptive statistics and categorical variables will be summarized with frequencies and percentages. The summaries will be provided for each analysis set and OA phenotype. Subject level listings will also be provided.

10.3. Medical History

Medical history will be collected at screening and reassessed at the Week 24 (EOS) / Early Termination visit to record any changes. A summary of reported medical history will be provided by system category for each treatment group in the SAS. A subject-level listing will provide further information on each event.

11. PROTOCOL DEVIATIONS

A protocol deviation is defined as any change, divergence or departure from the study design or procedures of a research protocol that is under the Investigator's control and that has not been approved by the site's Institutional Review Board (IRB).

Deviations are summarized into one of the following categories:

- Informed Consent
- Enrollment
- Procedures
- Labs/Specimens
- Study Visits
- Source Documentation
- Investigational Product
- Diaries, Questionnaires, or Patient Reported Outcomes
- IXRS (Interactive Response Technologies)
- Subject Non-Compliance
- Adverse Events

Deviations are categorized as major or minor by a cross-functional team according to pre-defined criteria established in the Protocol Deviation Classification Guideline.

- A major deviation is defined as a divergence from the protocol that materially (a) reduces the quality or completeness of the data, (b) makes the informed consent inaccurate, or (c) impacts a subject's safety, rights or welfare.
- A minor deviation is defined as a divergence from the protocol that deviates from the procedures and guidelines outlined in the protocol, but is not classified as a major deviation (i.e. the deviation does not materially (a) reduce the quality or completeness of the data, (b) make the informed consent inaccurate, or (c) impact a subject's safety, rights or welfare).

Protocol deviations will be summarized by site, category and classification, and listed for each subject.

12. ANALYSIS SOFTWARE

All data processing, summarization and analyses will utilize SAS® Version 9.4, equipped with Cytel Inc.'s PROC MCPMOD.

13. REFERENCES

Bang, H., Flaherty, S.P., Kolahi, J., Park, J. (2010). "Blinding assessment in clinical trials: A review of statistical methods and a proposal of blinding assessment protocol." Clinical Research and Regulatory Affairs **27**(2): 42-51.

Kincaid CD (2005). "Guidelines for Selecting the Covariance Structure in Mixed Model Analysis." Proceedings of the Thirtieth Annual SAS® Users Group International Conference. Paper 198-30.

KOOSscoring2012.pdf. (2012). Available at: http://www.koos.nu/KOOSscoring2012.pdf [Accessed 24 May 2018].

Pham, T., D. Van Der Heijde, M. Lassere, R. D. Altman, J. J. Anderson, N. Bellamy, M. Hochberg, L. Simon, V. Strand, T. Woodworth, M. Dougados and O. Omeract (2003). "Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria." J Rheumatol 30(7): 1648-1654.

Ting N., Chen D.-G., Ho S. and Cappelleri J. C. (2017). Phase II Clinical Development of New Drugs. Singapore, Springer Singapore.