Statistical Analysis Plan

A Randomized, Double-Blind, Placebo Controlled, Multi-Center Study of a Single Injection Cross-Linked Sodium Hyaluronate Combined with Triamcinolone Hexacetonide (Cingal®) to Provide Symptomatic Relief of Osteoarthritis of the Knee

Protocol Number: CINGAL 19-01 Version 1.0
Protocol Issue Date: 13 Dec2018
SAP Date:26AUG2020
SAP version number- V1.0
Anika Therapeutics, Inc.

Prepared by: Gary Stevens, Ph.D.
DynaStat Consulting, Inc.

Approved by:

	Gary Stevens, Ph.D.
Anika Therapeutics, Inc.	Dynastat Consulting, Inc.

Table of Contents

Contents

1.0 Introduction

This document presents the statistical analysis plan (SAP) for the Anika Therapeutics, Inc. protocol Cingal 19-01. This incorporates the data listings, summary tables, and statistical analyses. This document is based on the original final protocol and case report forms (CRFs).

2.0 Study Design

This is a multi-center, randomized, double-blind, parallel arm, placebo-controlled trial to determine the contribution of Triamcinolone Hexacetonide (TH) to pain relief, both in terms of magnitude and duration, when used within a single injection of Cingal® compared to a single injection of TH in subjects with Osteoarthritis (OA) of the knee.

Subjects with OA defined as Kellgren-Lawrence (K-L) II or III in the index knee will be eligible for this study [9, 10]. Felson et al 2004, encouraged the use of both OA symptoms and radiographic changes in the assessment of OA [11]. This study will employ both methods to screen subjects. Structural severity will be evaluated with the K-L classification score, a composite index of the presence and severity of joint space narrowing, osteophytes, sclerosis, deformity and cysts.

For the evaluation of symptomatic severity, two main domains are important [12]. The first is pain and the second is functional impairment. Other domains often used include subject's overall assessment, ROM and performance. Domains identified by OMERACT as core variables to be used in clinical trials involving OA are pain, function and the Patient Global Assessment which will be captured as part of this study [12].

Baseline and post-treatment pain, physical function and stiffness will be measured using the WOMAC questionnaire. VAS and NRS pain, ROM, Patient and Evaluator Global Assessment and the EuroQoL-5D-5L will be used to assess symptomatic severity throughout the study. In addition, the number of acetaminophen pills taken will be captured as an indirect measure of pain and will be done at follow up visits.

Subjects meeting the inclusion/exclusion criteria will be randomized to receive a single injection of Cingal®, Triamcinolone Hexacetonide (TH) or placebo in the index knee. Since there is a difference in volume between Cingal®, TH and placebo, the treating physician will not be considered blinded. To maintain the double-blind design of the study, there will be a person assigned to the role of Treating Physician and one person assigned to the role of Blinded Evaluator. The Treating Physician, most often the PI, will administer the injection but will not participate in the evaluation of study treatment effectiveness. A second individual, designated as the Blinded Evaluator, is blinded to treatment and will complete the pre- and post-treatment Evaluator Global Assessment, knee exams and ROM measurements. To maintain the Subject blinding, the injection syringe will be prepared separate from the patient and the injection will be masked from the subject.

The Subject will be trained on how to complete the WOMAC, VAS Pain, NRS Pain, Patient Global Assessment and EuroQoL.

The Blinded Evaluator will collect and record AEs from the subjects and consult the Treating Physician only as needed in the management of AEs. The Blinded Evaluator will be a physician, research nurse, registered physiotherapist or physician assistant trained to perform the assessments outlined in the protocol.

Up to 40 sites in US may participate in the study to enroll and treat 231 subjects. Subject participation will last approximately 7 months with follow-up visits scheduled at weeks 1, 3, 6, 12, 18, and 26 after the treatment injection.

3.0 Objectives

To determine the contribution of Triamcinolone Hexacetonide (TH) to pain relief, both in terms of magnitude and duration, when used within a single injection of Cingal® compared to a single injection of TH in subjects with Osteoarthritis (OA) of the knee. A saline placebo is included within the trial to set the expectation of a return to pain in Subjects.

3.1 PRIMARY EFFICACY ENDPOINT

- The change from Baseline in knee pain as measured by the WOMAC Pain Score (100 mm VAS) at 26 weeks post treatment comparing the Cingal® arm to the TH arm, the Cingal arm to the placebo arm, and the TH arm to the placebo arm.

3.2 SECONDARY EFFICACY ENDPOINTS

- The responder rate as identified by the Outcomes Measures for Rheumatic Arthritis Clinical Trials-Osteoarthritis Research Society International (OMERACT-OARSI) responder index at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in knee pain as measured by daily VAS Pain Score and average weekly VAS Pain Score (100 mm VAS) through 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in knee pain as measured by daily NRS Pain Score and average weekly NRS Pain Score (0-10 NRS) through 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in WOMAC Physical Function score at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.

- The change from Baseline in WOMAC Stiffness score at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in Total WOMAC score at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in the Patient Global Assessment at 26 weeks post treatment in the Cingal® arm compared to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in the Evaluator Global Assessment at 26 weeks post treatment in the Cingal® arm compared to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The usage of rescue medication through 26 weeks post treatment in the Cingal® arm compared to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.

3.3 EXPLORATORY ENDPOINTS

Any comparisons between arms (Cingal®, TH, Placebo), within arms and / or time points (from baseline through to 26 weeks) not described in the primary or secondary endpoints may be presented in the exploratory endpoints including:

- EuroQoL (EQ-5D-5L)
- WOMAC Pain Score (100mm VAS)
- Daily VAS pain scores & weekly average VAS pain scores
- Daily NRS pain scores & weekly average NRS pain scores
- OMERACT-OARSI
- Total WOMAC
- WOMAC Stiffness Score
- WOMAC Physical Function
- Patient Global Assessment
- Evaluator Global Assessment
- Range of Motion
- Rescue Medication Use
- Number of Treatment Failures

3.4 SAFETY ENDPOINT:

The incidence, timing, severity, and relationship to treatment of all Adverse Events (AE) will be collected and coded using Medical Dictionary for Regulatory Activities (MedDRA). Local injection site and non-local events will be recorded separately.

4.0 Statistical Methods in Protocol

4.1 STATISTICAL METHODS

This Statistical Analysis Plan provides the details of the statistical analysis of the study data that is described in the protocol.

4.1.1 Sample Size

This is a pilot study designed to estimate as the primary assessment, the differences between Cingal and Triamcinolone Hexacetonide (TH) with respect to the difference in the change in WOMAC pain score from baseline to 26 weeks post treatment. For convenience the sample size will be assessed in a 3:3:1 ratio (Cingal: TH: Placebo) with the sample sizes being 99:99:33, for a total sample size of 231 patients treated. It should be noted that if the standard deviation for the difference in change from baseline in the WOAMC Pain Score between Cingal and TH is approximately 20 points and a difference of 8 points would yield 80% power of detecting that difference. If the difference is 10 points, this sample size yields over 90% power of detecting that difference.

Since the primary analysis in this study is an intent to treat analysis and the missing data for a patient will be imputed using mixed effects models, no dropout rate will be assumed for this study. Thus, the total sample size will be 231 patients treated.

4.1.2 Primary Efficacy Endpoint

4.1.2.1 The primary efficacy endpoint is the change from baseline in knee pain as measured by the WOMAC Pain Score (1 00 mm VAS) at 26 weeks post treatment comparing the Cingal Arm to the triamcinolone hexacetonide (TH) Arm, the Cingal Arm to the Placebo Arm, and the TH Arm to the Placebo Arm .

Data will be analyzed via an analysis of variance (ANOVA) with a term for treatment and baseline pain score as a covariate. The primary hypothesis(es) will be tested by a predefined contrast for comparing the Cingal Arm to the triamcinolone hexacetonide (TH) Arm, the Cingal Arm to the Placebo Arm, and the TH Arm to the Placebo Arm. Formally, the hypotheses to be tested is:

H_0: $\mu_{DC} = \mu_{DT}$ versus H_A: $\mu_{DC} \neq \mu_{DT}$. (Primary Cingal to TH)
H_0: $\mu_{DC} = \mu_{DP}$ versus H_A: $\mu_{DC} \neq \mu_{DP}$. (Primary Cingal to Placebo)
H_0: $\mu_{DT} = \mu_{DP}$ versus H_A: $\mu_{DT} \neq \mu_{DP}$. (Primary TH to Placebo)

Where μ_{DC} is the mean change from baseline in the WOMAC Pain Score for the Cingal Arm at 26 weeks, μ_{DT} is the mean change from baseline in the WOMAC Pain Score for the triamcinolone hexacetonide Arm at 26 weeks, and μ_{DP} is the mean change from baseline in the WOMAC Pain Score for the Placebo Arm at 26 weeks. Since there are three treatment groups in this study, this hypothesis will be tested using a one way ANOVA and constructing a contrast for this hypothesis.

This process is the same as using a baseline pain adjusted two-sample t-test with the common variance estimate from the ANOVA.

The anticipated SAS code for this analysis is given as:
The mean difference will be estimated and tested using the contrasts from PROC GLM on the responses at the 26 week time period.

```
PROC GLM DATA=EFF;
 WHERE VISITNO = 26;
 CLASS TRT;
 MODEL WOMDIFF = BASEWOM TRT / TYPE3 ALPHA=0.05;
 ESTIMATE 'CINGAL VS PLACEBO'  TRT 1 -1 0 ;
 ESTIMATE 'CINGAL VS TH'       TRT 1 0 -1 ;
 ESTIMATE 'PLACEBO VS TH'      TRT 0 1 -1 ;
 LSMEANS TRT / DIFF CL ALPHA=0.05;
FORMAT TRT TRTF. ;
RUN;
```

In this model, the WOMDIFF is the difference in WOMAC Pain Scores from baseline. The variable TRT is a categorical variable where 1 = Cingal, 2=Placebo, and 3=TH. The variable BASEWOM is the baseline WOAMC scores and the variable VISITNO is the visit number identifier. ODS statements will be used to output the contrasts and the mean differences into data sets that can be easily displayed using PROC REPORT. The primary analysis will be comparing Cingal to TH.

The tables for this analysis will be displayed as:

Summary of Least Squares Means at 26 weeks – WOMAC Pain Score

Treatment Group	**Mean Percent Improvement LS Mean**	**Standard Error**
Cingal		
Placebo		
TH		

Summary of Differences in Mean Responses at 26 Weeks – WOMAC Pain Score

Treatment Group	**Mean Percent Improvement**	**Standard Error**	**Lower Confidence Limit**	**Upper Confidence Limit**	**p-value**
Cingal – Placebo					
Cingal – TH (Primary)					
Placebo - TH					

4.1.3 Secondary Efficacy Endpoints

4.1.3.1 The responder rate as identified by the Outcomes Measures for Rheumatic Arthritis Clinical Trials-Osteoarthritis Research Society International (OMERACT-OARSI) responder index at 26 weeks post treatment comparing the Cingal® arm to the TH arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo Arm.

The response rate for the Cingal group and Triamcinolone Hexacetonide (TH) group will be calculated and tested using Fisher's exact test. Formally, the hypothesis to be tested is:

H0: $\pi_C = \pi_T$ versus H0: $\pi_C \neq \pi_T$

Where π_C is the responder rate for the Cingal group and π_T is the responder rate for the triamcinolone hexacetonide group.

The response rate for the Cingal Arm, the Placebo Arm and Triamcinolone Hexacetonide (TH) Arm will be calculated and tested using Fisher's exact test. Formally, the hypotheses to be tested are:

H0: $\pi_C = \pi_P$ versus H0: $\pi_C \neq \pi_P$
H0: $\pi_T = \pi_P$ versus H0: $\pi_T \neq \pi_P$

Where π_C is the responder rate for the Cingal Arm, π_T is the responder rate for the triamcinolone hexacetonide Arm, and π_P is the responder rate for the Placebo Arm.

This will be estimated and tested using PROC FREQ with the EXACT option. The results will be displayed similar to:

Summary of the OMERACT-OARSI Responder Rates at 26 Weeks

Parameter	Cingal	Placebo	TH
Estimate n/N (%)			
Cingal vs TH (Primary) (p-value)			
Cingal vs Placebo (p-value) Placebo vs TH (p-value)			

4.1.3.2 Other Secondary Endpoints .

All of the following endpoints will be analyzed using the same methodology as the primary endpoint:

- The change from Baseline in knee pain as measured by daily VAS Pain Score and average weekly VAS Pain Score (100 mm VAS) through 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in knee pain as measured by daily NRS Pain Score and average weekly NRS Pain Score (0-10 NRS) through 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in WOMAC Physical Function score at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in WOMAC Stiffness score at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in Total WOMAC score at 26 weeks post treatment comparing the Cingal® arm to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.
- The change from Baseline in the Patient Global Assessment at 26 weeks post treatment in the Cingal® arm compared to the TH arm and Placebo arm, the Cingal arm to the Placebo arm, and the TH arm to the Placebo arm.

That is, Data will be analyzed via an analysis of variance (ANOVA) with a term for treatment and baseline pain score as a covariate. The primary hypothesis(es) will be tested by a predefined contrast for comparing the Cingal Arm to the triamcinolone hexacetonide (TH) Arm, the Cingal Arm to the Placebo Arm, and the TH Arm to the Placebo Arm.. Formally, the hypotheses to be tested is:

H_0: $\mu_{DC} = \mu_{DT}$ versus H_A: $\mu_{DC} \neq \mu_{DT}$. (Primary Cingal to TH)
H_0: $\mu_{DC} = \mu_{DP}$ versus H_A: $\mu_{DC} \neq \mu_{DP}$. (Primary Cingal to Placebo)
H_0: $\mu_{DT} = \mu_{DP}$ versus H_A: $\mu_{DT} \neq \mu_{DP}$. (Primary TH to Placebo)

Where μ_{DC} is the mean change from baseline in the Endpoint Score for the Cingal Arm at 26 weeks, μ_{DT} is the mean change from baseline in the Endpoint Score for the triamcinolone hexacetonide Arm at 26 weeks, and μ_{DP} is the mean change from baseline in the Endpoint Score

for the Placebo Arm at 26 weeks. Since there are three treatment groups in this study, this hypothesis will be tested using a one way ANOVA and constructing a contrast for this hypothesis. This process is the same as using a baseline Endpoint Score adjusted two-sample t-test with the common variance estimate from the ANOVA.

The anticipated SAS code for this analysis is given as:
The mean difference will be estimated and tested using the contrasts from PROC GLM on the responses at the 26 week time period.

```
PROC GLM DATA=EFF;
 WHERE VISITNO = 26;
 CLASS TRT;
 MODEL PARMDIFF = BASEPARM TRT / TYPE3 ALPHA=0.05;
 ESTIMATE 'CINGAL VS PLACEBO'  TRT 1 -1 0 ;
 ESTIMATE 'CINGAL VS TH'       TRT 1 0 -1 ;
 ESTIMATE 'PLACEBO VS TH'      TRT 0 1 -1 ;
 LSMEANS TRT / DIFF CL ALPHA=0.05;
FORMAT TRT TRTF. ;
RUN;
```

In this model, the PARMDIFF is the difference in Endpoint Score from baseline. The variable TRT is a categorical variable where 1 = Cingal, 2=Placebo, and 3=TH. The variable BASEPARM is the baseline Endpoint Score and the variable VISITNO is the visit number identifier. ODS statements will be used to output the contrasts and the mean differences into data sets that can be easily displayed using PROC REPORT. The primary analysis will be comparing Cingal to TH.

The tables for this analysis will be displayed as:

Summary of Least Squares Means at 26 weeks – Endpoint Score

Treatment Group	**Mean Percent Improvement LS Mean**	**Standard Error**
Cingal		
Placebo		
TH		

Summary of Differences in Mean Responses at 26 Weeks – Endpoint Score

Treatment Group	**Mean Percent Improvement**	**Standard Error**	**Lower Confidence Limit**	**Upper Confidence Limit**	**p-value**
Cingal – Placebo					
Cingal – TH (Primary)					
Placebo - TH					

4.1.3.9 The usage of rescue medication through 26 weeks post treatment comparing the Cingal group to the triamcinolone hexacetonide group.

The percentage of subjects using rescue medications will be calculated and tested using Fisher's exact test. Formally, the hypothesis to be tested is:

H0: $\pi_C = \pi_T$ versus H0: $\pi_C \neq \pi_T$
H0: $\pi_C = \pi_P$ versus H0: $\pi_C \neq \pi_P$
H0: $\pi_T = \pi_P$ versus H0: $\pi_T \neq \pi_P$

Where π_C is the usage rate for the Cingal Arm, π_T is the usage rate for the triamcinolone hexacetonide Arm, and π_P is the usage rate for the Placebo Arm.

This will be estimated and tested using PROC FREQ with the EXACT option. The results will be displayed similar to:

Summary of the Usage Rates at 26 Weeks

Parameter	**Cingal**	**Placebo**	**TH**
Estimate n/N (%)			
Cingal vs TH (Primary) (p-value)			
Cingal vs Placebo (p-value) Placebo vs TH (p-value)			

4.1.4 Exploratory Endpoints

Any comparisons between groups (Cingal, Placebo, TH), within groups and / or time points (from baseline through to 26 weeks) not described in the primary or secondary endpoints may be presented in the exploratory endpoints including:

- EuroQoL (EQ-5D)
- WOMAC Pain Score (100mm VAS)
- OMERACT-OARSI
- Total WOMAC
- WOMAC Stiffness Score
- WOMAC Physical Function Score
- Patient Global Assessment
- Evaluator Global Assessment
- Range of Motion
- Rescue Medication Usage: Difference in analgesic use measured by number of pills taken between visits will be compared between treatment groups descriptively.
- Number of Subjects Considered Treatment Failures: Treatment Failure: A subject who undergoes a procedure or uses a medication (other than the rescue medication) for the

treatment of OA in the index knee at any time after the study injection through the 12 week visit.

For the exploratory analyses, the continuous variables will be analyzed via an analysis of variance and contrasts used to assess the prescribed comparisons. For the discrete variables, a Fisher's exact test will be used to assess the desired comparisons.

4.2 STATISTICAL METHODS

Tabulation of summary statistics, graphical presentations, and statistical analyses will be performed using SAS® software 9.4 or higher version. Where not otherwise specified, the last pre-treatment observation will be used as baseline for calculating post-treatment changes from baseline. The primary presentations and analyses will be based on data pooled across study centers. Relevant summaries for individual centers, or combinations of centers, may be presented for primary data. All testing and confidence intervals will use a significance level of 5%.

4.3 DEMOGRAPHIC AND BASELINE CHARACTERISTICS

All demographic and baseline characteristics will be tabulated by treatment group and a test of homogeneity between the treatment groups will be conducted. For continuous variables (e.g. age, height, weight) a one way ANOVA will be used. For categorical variables (e.g. gender, race), a Fisher's exact test or chi-squared test will be used. Medical history findings, physical examinations and concomitant medications will be tabulated by treatment group.
The baseline Chemistry, Hematology, Urinalysis, Vital Signs, and BMI data will be summarized via descriptive statistics and tested for homogeneity using a one way ANOVA.

4.4 ADVERSE EVENTS

All AEs will be coded according to MedDRA. Safety assessments will include Treatment-Emergent Adverse Events (TEAEs) which are defined as AEs with an Investigator assessment of definitely, probably, or possibly related to CTM. TEAEs will be summarized with frequencies and percentages by system organ class and preferred term, severity, and relationship to study CTM for each treatment group. In summaries of TEAEs by severity and relationship to CTM for subjects reporting multiple episodes, all reported events will be included, not only the worst reported case. Serious Adverse Events will also be presented by relationship to the CTM.
The number of subjects with at least one AE will be tabulated for each treatment group. Differences between the treatment groups will be tested using Fisher's exact test. Then the number of AEs for each treatment group will also be tabulated.

The number of subjects and the number of AEs will be tabulated by severity, relationship, and local injection site specific events versus non local events.

4.5 SUBJECT POPULATIONS

All safety analyses will be conducted on all subjects who undergo treatment in any group.

The primary analysis on the primary endpoint will be performed on the ITT populations using the Multiple Imputation Methodology. The Multiple Imputation Methodology will use a mixed effects repeated measures model to predict the missing values. All Primary and Secondary endpoints will be analyzed using the ITT population.

It has been determined that any missing data in this study will follow the Missing at Random (MAR) assumption and is justified by the following section from the guide on missing data.

Missing Data Because of Attrition in the Course of the Study The longer the planned length of a clinical trial, the greater the chance that participants will drop out of the trial due to their moving out of the area or otherwise experiencing changes in their lives that preclude or complicate further participation. If dropping out due to these situations is known to be unrelated to changes in health status, an MAR assumption for the missing values seems justified;

Since none of these patients are likely to be in a life threatening disease situation and many have had knee pain for a period of time, the missing data will unlikely be due to any treatment effect or lack thereof. Previous studies have demonstrated a missing data rate of less than 1%. Thus, any missing data in this study follows that paradigm and thus should be considered MAR. Under this MAR assumption, the mixed effects repeated measures analysis yields unbiased estimates of the treatment effects and thus will be utilized in the analysis.

A secondary analysis will be conducted on the Per Protocol (PP) population. Since the primary endpoint is at 26 weeks, this is all subjects who complete the 26 week visit and who are not major violators of the protocol. For all other visits, this is defined as the subjects who complete those visits according to the protocol.

4.6 ADDITIONAL ANALYSES

All of the analyses are performed on the data with the specified covariates or other factors assumed in the model. If it is determined that certain factors may influence the outcomes of the endpoints, then additional analyses will be performed. For the continuous variables, the factors will be added to the ANOVA model and evaluated in a stepwise fashion for significance. For the discrete variables, the data will be analyzed via a GEE model with the factors added to the model.