Official Title: A Randomized, Placebo-Controlled, Double-Blind Pilot Study to Evaluate the Effect of GRF6021 on Intracellular Signaling Cascades in Blood Leukocytes and Postoperative Recovery Following Primary Hip or Knee Arthroplasty

NCT Number: NCT03981419

Document Date: SAP Version 1: 08-October-2020

STATISTICAL ANALYSIS PLAN

Based on:
Protocol Version 3.0, Dated 14NOV2019

SPONSOR:	Alkahest, Inc.
PROTOCOL NUMBER:	AKST6021-211
PROTOCOL TITLE:	A Randomized, Placebo-Controlled, Double-Blind Pilot Study to Evaluate the Effect of GRF6021 on Intracellular Signaling Cascades in Blood Leukocytes and Postoperative Recovery Following Primary Hip or Knee Arthroplasty
ORIGINAL PROTOCOL DATE:	11MAR2019

Created by: [redacted]

SAP Approval

By signing the following, I agree to the contents in the Statistical Analysis Plan and its associated attachments. Once the SAP has been signed, the analyses and programming of the TFLs based upon this document can proceed. Any modifications to the SAP and TFLs made after signing may result in a change order.

Approved by:

Name	Title	Signature	Date
[redacted], MD, PhD	[redacted], Clinical Development	DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this document Signing Time: 08-Oct-2020 | 8:05:44 AM PDT D45401DB10A04F25A1F5F77BBDD81AEC	08-Oct-2020
[redacted], MPH	[redacted]	DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this document Signing Time: 08-Oct-2020 | 9:05:29 AM PDT 852FBCE678654B2E89387C335E98EDEB	8-Oct-2020

Abbreviations

Abbreviation	Definition
3D-CAM	3-Minute Diagnostic Interview for Confusion
ADLB	Laboratory Analysis Dataset
AE	Adverse Event
AESI	Adverse Event of Special Interest
ALT	Alanine aminotransferase
ASA	American Society of Anesthesiologists
AST	Aspartate aminotransferase
AUC	Area Under the Curve
BDI-II	Beck Depression Index-II
CBC	Complete blood count
CSR	Clinical Study Report
CyTOF	Mass cytometry
CytoF	Mass cytometry
EBL	Estimated blood loss
ECG	Electrocardiogram
GGT	Gamma-glutamyl transferase
HIV	Human immunodeficiency virus
HR	Heart rate
Iba1	Ionized calcium binding adaptor molecule 1
IgA	Immunoglobulin A
ITT	Intent-to-treat
IV	Intravenous
KM	Kaplan-Meier
MedDRA	Medical Dictionary for Regulatory Activities
PACU	Post-anesthesia care unit
POMS	Profile of Mood States
PSS	Perceived Stress Scale
PT	Preferred Term
PTT	Partial thromboplastin time
RBC	Red blood cell
RR	Respiration rate
SAE	Serious adverse event
SAP	Statistical Analytical Plan
SF-36	Short Form-36
SOC	System Organ Class
SRS	Surgery Recovery Scale
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

Table of Contents

STATISTICAL ANALYSIS PLAN **1**

SAP APPROVAL **2**

ABBREVIATIONS **4**

TABLE OF CONTENTS **5**

1 INTRODUCTION **7**

2 OBJECTIVES **7**

2.1 PRIMARY OBJECTIVE 7
2.2 SECONDARY OBJECTIVES 7

3 ENDPOINTS **7**

3.1 PRIMARY ENDPOINT 7
3.2 SECONDARY ENDPOINTS 7

4 STUDY OVERVIEW **8**

4.1 STUDY DESIGN 8
4.2 TREATMENT ASSIGNMENT AND DOSE REGIMEN 8
4.3 STATISTICAL HYPOTHESIS 9
4.4 SAMPLE SIZE JUSTIFICATION 9
4.5 RANDOMIZATION/UNBLINDING 9
4.5.1 Randomization *9*
4.5.2 Unblinding *9*
4.6 INTERIM ANALYSES 9
4.7 STUDY ASSESSMENT TIME POINTS 9
4.8 SCHEDULE OF EVENTS 11

5 STATISTICAL METHODOLOGY **14**

5.1 GENERAL CONSIDERATIONS 14
5.2 STUDY POPULATIONS FOR ANALYSIS 14
5.2.1 Intention-to-Treat (ITT) Population *14*
5.2.2 Safety Population *14*
5.2.3 Evaluable Population *15*
5.2.4 Per-Protocol (PP) Population *15*
5.2.5 Subgroup Analyses *15*
5.3 SUBJECT DISPOSITION 15
5.4 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 15
5.4.1 Demographics *15*
5.4.2 Baseline Characteristics *16*
5.4.3 Medical History *16*
5.4.4 Tobacco and Substance Use *16*
5.5 PROTOCOL DEVIATIONS 16
5.6 METHODS FOR HANDLING MISSING DATA 16
5.7 EFFICACY ANALYSES 16
5.7.1 Primary Endpoint *16*
5.7.2 Secondary Efficacy Endpoints *16*
5.7.2.1 Time to 50% recovery of the baseline value in the SRS 17
5.7.2.2 Time to a score of < 12/40 on a subset of questions from WOMAC, Pain Subscale. 17
5.7.2.3 Time to a score of < 18/60 on a subset of questions from the WOMAC, Physical Function Subscale. 18
5.7.2.4 End of Study Treatment Comparison of WOMAC 18
5.7.2.5 Change from Baseline in 3D-CAM 18
5.7.2.6 Change from baseline in the Short Form-36 (SF-36) 19

5.7.2.7 Change from baseline in the Beck Depression Inventory-II (BDI-II) 19
5.7.2.8 Opioid analgesic consumption during hospital stay and after discharge to end of study 20
5.7.2.9 Time to discharge 20
5.7.2.10 Perioperative outcomes 20
5.7.2.11 Change in functional status using the ActiGraph wearable device providing measurements for physical activity/function and sleep 21
5.7.2.12 The effects of GRF6021 on plasma proteomics 21
5.8 SAFETY ANALYSES 21
5.8.1 Extent of Exposure 21
5.8.2 Adverse Events 21
5.8.2.1 Summary for All TEAEs 22
5.8.2.2 Summary for serious TEAEs 22
5.8.2.3 Summary for AESI 22
5.8.2.4 Deaths 23
5.8.3 Laboratory Data 23
5.8.4 Vital Signs 24
5.8.5 12-lead ECG 25
5.8.6 Targeted Physical Examination 25
5.8.7 Pregnancy tests 25
5.8.8 Prior and Concomitant Medications 25
6 DOCUMENT VERSION CONTROL 26
REFERENCE 26
APPENDIX A - PROGRAMMING SPECIFICATIONS FOR TABLES AND LISTINGS 27

1 Introduction

This Statistical Analysis Plan (SAP) describes the planned statistical analyses to be performed for data from Protocol AKST6021-211, "A Randomized, Placebo-Controlled, Double-Blind Pilot Study to Evaluate the Effect of GRF6021 on Intracellular Signaling Cascades in Blood Leukocytes and Postoperative Recovery Following Primary Hip or Knee Arthroplasty". This SAP was created using Clinical Protocol AKST6021-211 Version 2.0 dated 15APR2019 and updated as necessary using Version 3.0 dated 14Nov2019.

The table and listing shells will be provided in separate files as attachments to this SAP.

2 Objectives

2.1 Primary Objective

To investigate the effect of GRF6021 on intracellular signaling cascades in blood leukocytes as determined by CyTOF

2.2 Secondary Objectives

- To investigate GRF6021 on clinical recovery parameters in subjects undergoing primary hip or knee arthroplasty
- To assess the safety and tolerability of GRF6021
- To assess the effects of GRF6021 on plasma proteomics

3 Endpoints

3.1 Primary Endpoint

Effect of GRF6021 on intracellular signaling cascades in blood leukocytes as determined by CyTOF

3.2 Secondary Endpoints

- Change from baseline in 3D-CAM.
- Time to 50% recovery of the baseline value in the SRS.
- End of study treatment comparison of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).
- Time to a score of < 12/40 on the WOMAC, Pain Subscale.
- Time to a score of < 18/60 on the WOMAC, Physical Function Subscale.
- Change from baseline in the SF-36.
- Change from baseline in the BDI-II.
- Opioid analgesic consumption during hospital stay and after discharge to end of study.
- Time to discharge.
- Perioperative outcomes (e.g., surgery duration, anesthesia duration, PACU stay duration, ASA status, EBL, type and amount of intraoperative fluids administered, type and amount of intraoperative blood products, dose and type of anesthesia (general, regional), and intraoperative opioid dose).

- Change in functional status using the ActiGraph wearable device providing measurements for physical activity/function and sleep (starting at least 3 days prior to surgery and continuing until the end of the study).
- Change from baseline in clinical laboratory parameters.
- Change from baseline in vital sign measurements.
- Change from baseline in ECG.
- The effects of GRF6021 on plasma proteomics.

4 Study Overview

4.1 Study Design

This is a randomized, placebo-controlled, double-blind, phase 2 pilot study conducted at a single site, to investigate the effects of GRF6021, a 5% human plasma protein fraction administered by intravenous (IV) infusion, on intracellular signaling cascades in blood leukocytes in subjects aged between 50 and 87 years undergoing primary hip or knee arthroplasty.

Approximately 45 subjects will be enrolled with the aim of having 40 evaluable subjects randomized in a 1:1 ratio to active treatment or placebo. Each subject will receive 4 infusions of 250 mL of active GRF6021 or 4 infusions of 250 mL of placebo:
1) 1 infusion on the day before surgery
2) 1 infusion on the day of surgery within 4 hours before surgery start (first incision)
3) 1 infusion on the day of surgery upon arrival in the postoperative care unit (within 5 hours after the first incision)
4) 1 infusion on the day after surgery

CyTOF, a multiplexed, high-content immune profiling technology, will be utilized to provide high-resolution surveillance of circulating immune cells and the response to GRF6021 infusions on surgical recovery. Additional functional and quality of life assessments will be measured at baseline and for 6 weeks after hospital discharge. A detailed overview of the timing of study assessments and interventions is presented in the Schedule of Events section below.

4.2 Treatment Assignment and Dose Regimen

A subject number will be assigned once the subject has signed the informed consent form. Enrolled subjects that have met all eligibility criteria will be randomized to GRF6021 (Group 1: active dosing) or placebo (Group 2: placebo dosing). Subjects will receive 4 infusions of 250 mL per the following schedule:
1) On the day before surgery
2) On the day of surgery within 4 hours before surgery start (first incision)
3) On the day of surgery upon arrival in the postoperative care unit (within 5 hours after the first incision)
4) On the day after surgery.

4.3 Statistical Hypothesis

The hypothesis of this study is that active treatment will be associated with immunomodulatory changes that have previously been associated with improved clinical recovery. However, this study is not designed to detect statistically significant differences between active and placebo on the primary endpoint because this is an exploratory study.

4.4 Sample Size Justification

A total of approximately 45 subjects will be enrolled with the aim of having 40 evaluable subjects randomized in a 1:1 ratio to active treatment or placebo. The primary endpoint of this study (differences in immunological and/or proteomic signatures among subject groups), the principal computational approach (machine learning) used, and the exploratory nature of the study prevent a traditional power analysis.

4.5 Randomization/Unblinding

This is a randomized and double-blind study.

4.5.1 Randomization

A randomization list, permuted blocks and stratified by sex (Male and Female) and surgical type (Hip and Knee) was generated by [redacted]

4.5.2 Unblinding

The infusion procedure of active and placebo agents will be identical to maintain the blind. The following measures will be taken to ensure adequate blinding during infusions: blinding of subjects, study coordinators, physicians, and raters to treatment allocation; use of unblinded infusion nurses; and measures to block view of the vial and drip chamber throughout the infusion. All blinded study site personnel (e.g. investigators, study coordinators and other blinded site staff), as well as the personnel involved in the blinded monitoring or conduct of the study, will be blinded to the individual subject treatment assignments. The blinded code for the trial will be broken only after all subject data has been recorded and verified and the database locked.

4.6 Interim Analyses

Safety will be monitored on an ongoing basis, and a Safety Evaluation Meeting may be triggered. Any such ad hoc interim safety analysis performed will be outside the scope of this SAP.

4.7 Study Assessment Time Points

The study consists of 15 protocol-specified visits which will be assessed as nominal visits from an analysis perspective:

1) Visit 1 Screening (Day -42 to -3)
2) Visit 2 Baseline (Days -7 to -3): Subjects will be contacted by phone to obtain the remote Baseline Visit assessments. Subjects will be randomized after eligibility has been confirmed.
3) Visits 3 through 5, Treatment (Days 1-3): on-site Baseline assessments will be gathered prior to the first infusion on Day 1. After the first infusion on Day 1, the subject will

be discharged from the facility and come back the following morning (Day 2) for the inpatient period. The inpatient treatment period is 2 days during which each subject will receive 3 additional doses of 250 mL of their assigned treatment allocation.

4) Visits 6 through 14 (Day 4 through Day 37±2): these are the follow-up visits, which will be conducted remotely by phone or electronically.
5) Visit 15, End of Study (Day 44 ± 2): the End of Study Visit will be conducted remotely by phone and/or electronically.
6) Early termination: If a subject has received at least one infusion but is terminated or terminates from the study early, the site should try to perform all assessments scheduled at the End of Study Visit (Visit 15) and complete the exit lab panel if the subject is still at the site and available and willing.

4.8 Schedule of Events

	Screening (On-site)	Baseline Visit (Remote)	Baseline Visit (On-site)/ Treatment (On-site)				Follow-up/End of Study (EOS) (Remote)									
Visit	1	2	3	4[a]	4[b]	5	6	7	8	9	10	11	12	13	14	15
Infusion Number			1	2	3	4										
Postoperative day			-1	0	0	1	2	4	7	10	14	17	21	28	35	42
Day	Day -42 to -3	Day - 7 to -3	1	2	2	3	4	6 ± 1	9 ± 1	12 ± 1	16 ± 1	19 ± 1	23 ± 2	30 ± 2	37 ± 2	44 ± 2
Week	-6 to -1	-1	1				1	1	1	2	2	3	3	4	5	6 EOS
Informed Consent	X															
Medical History	X															
Demographics	X															
Vital Signs	X		X[1,2]	X[1,2]	X[1,2]	X[1,2]										
Physical Exam/Height and Weight	X															
12-lead ECG[i]			X[1]			X[1,3]										
Randomization			X[1]													
Targeted Physical Exam			X[1]			X[1]										
Concomitant Medication Review	X	X	X[1]	X[1]		X[1]	X	X	X	X	X	X	X	X	X	X
Adverse Event Review		X	X[1]	X[1]		X[1]	X	X	X	X	X	X	X	X	X	X
GRF6021/Placebo Infusion			**X**	**X**	**X**	**X**										
Blood/Urine Sampling																
Screening Lab Panel[c]	X															
Exit Lab Panel[d]						X										

	Screening (On-site)	Baseline Visit (Remote)	Baseline Visit (On-site)/ Treatment (On-site)				Follow-up/End of Study (EOS) (Remote)									
Visit	1	2	3	4[a]	4[b]	5	6	7	8	9	10	11	12	13	14	15
Infusion Number			1	2	3	4										
Postoperative day			-1	0	0	1	2	4	7	10	14	17	21	28	35	42
Day	Day -42 to -3	Day - 7 to -3	1	2	2	3	4	6 ± 1	9 ± 1	12 ± 1	16 ± 1	19 ± 1	23 ± 2	30 ± 2	37 ± 2	44 ± 2
Week	-6 to -1	-1	1				1	1	1	2	2	3	3	4	5	6 EOS
Pregnancy Testing[e]	X		X[1]			X										
Proteomic Assessments[f]			X[1]	X[1]	X[1]	X[1]										
Efficacy Assessments																
CyTOF			X[1]	X[1]	X[1]	X[1]										
3D-CAM			X[1]		X	X[1]										
SRS		X	X[1]			X	X	X	X	X	X	X	X	X	X	X
Full WOMAC		X														X
Subset of WOMAC questions			X[1]			X[1]	X	X	X	X	X	X	X	X	X	X
SF-36		X														X
BDI-II		X														X
POMS		X														
PSS		X														
Opioid Consumption Assessment[g]	X				X	X	X	X	X	X	X	X	X	X	X	X
Perioperative Assessments					X											
ActiGraph[h]	(Provide Device)	(Place Device)	X			X	X	X	X	X	X	X	X	X	X	X

Note:
X^1: To be performed prior to infusion start (on Day 2, this is prior to the start of the first infusion).
X^2: To be performed:
Prior to each infusion, within 1 hour of infusion start: BP, HR, RR, and body temperature.
After the start of each infusion, at 15, 30, and 45 minutes: BP and HR.
At 60 minutes after infusion start: BP, HR, RR, and body temperature.
X^3: To be performed following final infusion on Day3.
a: Preoperative infusion on the day of surgery.
b: Postoperative infusion on the day of surgery.
c: Screening lab panel: Alkaline phosphatase, ALT, AST, bicarbonate, bilirubin (total), ionized calcium, chloride, creatinine, GGT, glucose (random), lactate dehydrogenase, magnesium, phosphate, potassium, protein total, sodium, blood urea nitrogen, albumin, PTT, PT/INR. CBC with differential (basophils, eosinophils, lymphocytes, monocytes, neutrophils). Urinalysis: blood, glucose, ketone, protein, pH, specific gravity, nitrite, leukocytes, urine creatinine, urine protein/creatinine ratio, (and reflex urine microscopy if indicated). Infectious serology (anti-HIV antibodies, anti-hepatitis B surface antigen, anti-hepatitis B core antigen, anti-hepatitis C antibodies), direct antiglobulin test, RBC antibody screen (including direct Coombs and anti-RBC antibody), serum IgA, and haptoglobin. Urine drug screen: cannabinoids, benzodiazepine, barbiturates, opiates, cocaine, amphetamines, phencyclidine.
d: Exit lab panel performed after the last infusion (Day 3): Infectious serology (anti-HIV antibodies, anti-hepatitis B surface antigen, anti-hepatitis B core antigen, anti-hepatitis C antibodies), direct antiglobulin test, RBC antibody screen (including direct Coombs and anti-RBC antibody), ALT, AST, bilirubin (total), ionized calcium, chloride, creatinine, potassium, protein total, sodium, blood urea nitrogen, PTT, PT/INR. CBC with differential (basophils, eosinophils, lymphocytes, monocytes, neutrophils). Urinalysis: blood, glucose, ketone, protein, pH, specific gravity, nitrite, leukocytes, urine creatinine, urine protein, urine protein/creatinine ratio, and reflex urine microscopy if indicated.
e: Serum pregnancy test required at screening. Serum and/or urine pregnancy test may be used at subsequent timepoints.
f: For proteomic assessments, fasting samples are preferred.
g: Baseline opioid consumption assessment will be included at screening during medication history. Subsequent opioid measurements will begin following surgery on Day 2 (Visit 4b).
h: The ActiGraph wearable device will be given to the subject at the screening visit. Between -7 and -3 days before surgery, the subject will be told by phone to put the ActiGraph on and wear it until admission. The ActiGraph will be removed for the surgery on Day 2 (Visits 4a and 4b).
i: On Day 3, second 12-lead ECG to be performed within +1 day

5 Statistical Methodology

5.1 General Considerations

This section presents the statistical approaches that are anticipated for the analysis of the study data. These approaches may at times require modifications due to unanticipated features of the data. Deviations from analyses summarized in this document will be noted in the CSR.

All statistical analyses will be performed using SAS® Version 9.4 or higher, unless otherwise noted. Data will be summarized using descriptive statistics only. Number of subjects (n), mean, median, standard deviation, minimum, and maximum will be summarized for continuous variables. Categorical variables will be summarized using frequencies and percentages. Time to event variables will be summarized using median, Q1 (25^{th} percentile) and Q3 (75^{th} percentile) from the Kaplan-Meier (KM) estimate of survival. Presentations will be by each treatment group and between-group differences, unless otherwise noted.

In general, all summary tables will be supported by a relevant subject data listing including all subjects who are randomized. The listings will include all data collected, and will be sorted by subject ID, and actual visit date, as applicable, unless otherwise noted.

Concomitant medications, intraoperative fluids, intraoperative blood products, and anesthesia will be coded by the WHO Drug Dictionary and adverse events and medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The dictionary versions will be documented in the Data Management Plan for the study.

Study day 1 is the first day of study treatment. Study day is the day relative to the study day 1.

Baseline is the last non-missing valid value prior to study treatment. Change from baseline is the post-treatment value minus baseline value.

5.2 Study Populations for Analysis

5.2.1 Intention-to-Treat (ITT) Population

All subjects who met the inclusion criteria and did not meet the exclusion criteria will be randomized and consist of the ITT population. All baseline characteristics will be summarized for the ITT population. Analysis will be done according to the treatment subjects were randomized to.

5.2.2 Safety Population

The safety population will consist of all subjects who received at least one dose of the study treatment. All safety endpoints will be summarized for the safety population. Analysis will be done according to the actual treatment subjects received.

5.2.3 Evaluable Population

All subjects who received all 4 doses will be included in the evaluable population. The evaluable population is the primary analysis population for the primary efficacy endpoint, and will also be used to summarize secondary efficacy endpoints. Analysis will be done according to the treatment that subjects were randomized to.

5.2.4 Per-Protocol (PP) Population

The PP population is the subset of evaluable population who received all 4 doses and completed the study without any protocol deviations that impact analyses. Such protocol deviations are departures from the approved protocol relating to the conduct of the study which may significantly impact the completeness, accuracy, and/or reliability of the study data or that may significantly affect the rights, safety or wellbeing of study participants. A listing of study subjects with protocol deviations which would preclude inclusion in the PP population will be provided by the Sponsor prior to database lock.

The PP population is a supportive analysis population for the primary efficacy endpoint, and will also be used to summarize secondary efficacy endpoints. Analysis will be done according to the treatment that subjects were randomized to.

5.2.5 Subgroup Analyses

Not applicable.

5.3 Subject Disposition

All subjects who signed the informed consent will be accounted for in Subject Disposition. The following disposition information will be summarized by the number and percentages of subjects:

- Subjects who signed the informed consent, and who comprise the ITT, Safety, Evaluable and PP populations
- Subjects who complete the study or withdraw prematurely along with the reason for discontinuation

A listing of screen failure subjects, including age, sex, race, ethnicity, and the inclusion/exclusion criteria not met, will be provided.

A listing of ITT subjects excluded from the PP population, including population flags (Safety/Evaluable) and the reason for exclusion, will be provided.

5.4 Demographics and other Baseline Characteristics

5.4.1 Demographics

Sex, age (in years), race, ethnicity, and baseline BMI (kg/m^2) will be summarized by treatment group for the ITT and Evaluable populations. BMI (kg/m^2) will be calculated as: 10,000*weight (kg) / the square of height (cm). If height is collected in inch, it will be converted to cm by dividing the value by 2.54. If weight is collected in pound, it will be converted to kg by dividing the value by 2.205.

5.4.2 Baseline Characteristics

Baseline characteristics will be summarized by treatment group for the ITT and Evaluable populations for the following:

- Surgical type (Knee or Hip)
- Short form of the Profile of Mood States (POMS-SF): calculated total score will be used
- Perceived Stress Scale (PSS): scores will be obtained by reversing responses (e.g., 0=4, 1=3, 2=2, 3=1, and 4=0) to the 4 positively stated items (items 4, 5, 7, and 8) and then summing across all scale items

5.4.3 Medical History

The number and percentage of subjects with any medical history will be summarized for the ITT population by treatment group, system organ class (SOC) and preferred terms (PT) using MedDRA dictionary.

5.4.4 Tobacco and Substance Use

Tobacco and substance use will not be summarized but will be provided as a listing.

5.5 Protocol Deviations

Protocol deviations will be summarized by deviation category and treatment group for the ITT, safety and evaluable populations.

5.6 Methods for Handling Missing Data

No imputation of missing data will be performed.

5.7 Efficacy Analyses

Efficacy analysis will evaluate the immune features and clinical response to study treatment. Except where indicated, efficacy parameters will be summarized by treatment group using both evaluable and PP populations.

5.7.1 Primary Endpoint

The analysis and summary of the primary endpoint will be performed by the Stanford group and is out of the scope of this SAP.

5.7.2 Secondary Efficacy Endpoints

All secondary efficacy endpoints will be descriptively summarized for both the evaluable and PP populations, unless otherwise specified. Per the Protocol, the following clinical outcomes are most important: Surgery Recovery Scale (SRS) (time to 50% recovery of the baseline score), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain (time to mild pain as indicated by a score <12/40), and WOMAC Physical Function (time to mild functional impairment as indicated by a score <18/60). All secondary endpoints will be summarized, as described below.

5.7.2.1 Time to 50% recovery of the baseline value in the SRS

The SRS[1] scale contains 13 items. The first 8 items each were scored from 1 to 6 (1 = not at all). The rest were scored from 1 to 5 (1 = not at all). For postoperative visits (Visit 5 through 8), the 'not applicable' responses will be coded into the same category as affirmations of 'not at all'. For all other time-points (Visit 2, Visit 3, Visit 9 through 15), 'not applicable' responses will be coded as missing data. SRS will be assessed at Visit 2 Baseline, Visit 3, and Visits 5 through 15.

SRS scores are obtained by reversing responses (e.g., 1=6, 2=5, 3=4, 4=3, 5=2 and 6 = 1) to the 5 negatively stated items (items 2, 4, 5, 6, and 7) and then summing across all scale items for an individual at baseline and each scheduled post-baseline time point. Higher score indicates better recovery. The KM estimate of survival will be used to summarize the time (in Days) from the date of randomization to 50% recovery of the baseline value in the SRS by treatment group using median, Q1 and Q3. Since SRS will only be evaluated at a clinic or remote visit, data is interval-censored. Subjects who have never achieved a 50% recovery of the baseline value until study exit, or who withdrew early prior to achieving a 50% recovery of the baseline value, will be censored at the last available SRS assessment day. If no SRS assessment is available, subjects will be censored at the date of randomization, whichever is later. The cumulative score (AUC) will also be presented by treatment group.

A KM survival curve stratified by treatment group will be provided. A listing of SRS assessment, including the total score at each time point and whether or not a subject achieved a 50% recovery of the baseline value at each post-baseline time point, will also be provided.

5.7.2.2 Time to a score of < 12/40 on a subset of questions from WOMAC, Pain Subscale.

The WOMAC Likert Scale version uses an 11-point scale anchored by the wording "no pain" and "extreme pain" for pain, and "no difficulty" and "extreme difficulty" for physical function. Higher scores on the WOMAC indicate worse pain and functional limitations. WOMAC will be assessed at Visit 2 Baseline, Visit 3, and Visit 5 through Visit 15.

Scores for each item of the pain subscale will be summed for an individual at baseline and each scheduled post-baseline time point. The KM estimate of survival will be used to summarize the time (in Days) from the date of randomization to a non-missing score of < 12 by treatment group using median, Q1 and Q3. Since WOMAC will only be evaluated at a clinic or remote visit, data is interval-censored. Subjects who have never achieved a non-missing score of < 12 until study exit, or who withdrew early prior to achieving a non-missing score of < 12, will be censored at the last available

WOMAC assessment day. If no WOMAC (Pain Subscale) assessment is available, subjects will be censored at the date of randomization, whichever is later. The cumulative score (AUC) will also be presented by treatment group.

A KM survival curve stratified by treatment group will be provided. A listing of WOMAC, Pain Subscale assessment, including the total pain score at each time point and whether or not a subject achieved a non-missing score of < 12 at each post-baseline time point, will also be provided.

5.7.2.3 Time to a score of < 18/60 on a subset of questions from the WOMAC, Physical Function Subscale.

Using the Likert Scale version of WOMAC, scores for each item of the physical function subscale will be summed for an individual at baseline and each scheduled post-baseline time point. The KM estimate of survival will be used to summarize the time (in Days) from the date of randomization to a non-missing score of < 18 by treatment groups using median, Q1 and Q3. Since WOMAC will only be evaluated at a clinic or remote visit, data is interval-censored. Subjects who have never achieved a non-missing score of < 18 at the end of the study, or who withdrew early prior to achieving a non-missing score of < 18, will be censored at the last available WOMAC (Physical Function Subscale) assessment day. If no WOMAC (Physical Function Subscale) assessment is available, subjects will be censored at the date of randomization, whichever is later. The cumulative score (AUC) will also be presented by treatment group.

A KM survival curve stratified by treatment group will be provided. A listing of WOMAC, Physical Function Subscale assessment, including the total physical function score at each time point and whether or not a subject achieved a non-missing score of < 18 at each post-baseline time point, will also be provided.

5.7.2.4 End of Study Treatment Comparison of WOMAC

Scores for each item of WOMAC will be summed for an individual at Visit 15/End of Study. The calculated score will be used to summarize the WOMAC score by treatment group.

5.7.2.5 Change from Baseline in 3D-CAM

Each item in the 3D-CAM instrument directly informs one of the 4 CAM features including acute onset of mental status change or fluctuating course of cognition, inattention, disorganized thinking, and altered level of consciousness. For all items on the assessment, the patient's answer is "incorrect," "correct," "no," or "yes." For all items, if the patient's answer is 'incorrect' or 'yes', then the appropriate (unshaded) column (representing a CAM feature) on the right side is checked. Delirium is considered present

if the following features are present: Feature 1) Acute onset or fluctuating course and Feature 2) Inattention and either Feature 3) Disorganized thinking or Feature 4) Altered level of consciousness. 3D-CAM will be assessed at Visit 3, Visit 4 postoperative, and Visit 5. Delirium will be determined according to The 3D CAM Training Manual for Research[2].

The number of subjects assessed via 3D-CAM, and the number and percentage of subjects with delirium present will be summarized by treatment group at baseline and each post-baseline scheduled time point. The difference from baseline in the number and percentage of subjects with delirium present will also be summarized at each post-baseline scheduled time point.

A listing of 3D-CAM assessments, including an evaluation of whether or not delirium is present, will be provided.

5.7.2.6 Change from baseline in the Short Form-36 (SF-36)

The SF-36 yields 8 sub-scale scores and 2 summary scores for mental and physical health. SF-36 will be assessed at Visit 2 Baseline and Visit 15/End of Study. The 8 sub-scale scores will be determined from the RAND Corporation web site and the mental and physical health scores will be derived using norm-based (standard) scoring.

The summary score for mental health andthe summary score for physical health will be summarized by treatment group at baseline and Visit 15/End of Study. The change from baseline in each of these scores at Visit 15/End of Study will be summarized.

A listing of SF-36 assessments, including the summary scores for mental and physical health and the total scores, will be provided.

5.7.2.7 Change from baseline in the Beck Depression Inventory-II (BDI-II)

The BDI-II 21-item self-administered survey is scored on a scale of 0–3 in a list of four statements arranged in increasing severity about a particular symptom of depression. BDI-II will be assessed at Visit 2 Baseline and Visit 15/End of Study.

The total scores will be summed for an individual and summarized by treatment group at baseline and Visit 15/End of Study. The change from baseline in the total scores at Visit 15/End of Study will be summarized.

5.7.2.8 Opioid analgesic consumption during hospital stay and after discharge to end of study

The total opioid consumption (mg*days) will be summarized by treatment group. The number and percentage of subjects will also be summarized by ATC level 3 terms and standardized medication names.

In addition, the KM estimate of survival will be used to summarize the time (in Days) from the date of randomization to cessation by treatment group using median, Q1 and Q3. Any Subject who do not stop using opioid at the end of the study will be censored at the study exit date. Subjects who are lost to follow-up, died, or who withdrew early prior to cessation, will be censored at the last known date of opioid consumption. If subjects never use opioid, they will be censored at the date of randomization.

5.7.2.9 Time to discharge

The KM estimate of survival will be used to summarize the time (in Days) from the date of randomization to discharge by treatment group using median, Q1 and Q3. Any Subject not discharged from the hospital at the end of the study will be censored at the study exit date. Subjects who are lost to follow-up, died, or who withdrew early prior to hospital discharge, will be censored at the last known date in the study. If subjects are not hospitalized, they will be censored at the date of randomization.

A KM survival curve stratified by treatment group will be provided. The hospital discharge date will be included in the subject disposition listing.

5.7.2.10 Perioperative outcomes

Perioperative outcomes include surgery duration, anesthesia duration, post-anesthesia care unit (PACU) stay duration, American Society of Anesthesiologists (ASA) status, estimated blood loss (mL), type and amount of intraoperative fluids administered, type and amount of intraoperative blood products, dose and type of anesthesia (general, regional), and intraoperative opioid dose.

5.7.2.10.1 Perioperative assessments

The surgery duration (in mins), anesthesia duration (inmins), PACU stay duration (in mins), estimated blood loss (mL), amount of intraoperative fluids administered will be summarized by treatment group. In addition, the number and percentage of subjects in each of the 5 ASA classes will be summarized by treatment group.

5.7.2.10.2 Intraoperative fluids administered

Intraoperative fluids administered will be summarized by ATC level 3 terms and standardized medication names. The amount of intraoperative fluids administrated (ml) will also be summarized.

5.7.2.10.3 Intraoperative blood products administered

Intraoperative blood products administered will be summarized by ATC level 3 terms and standardized medication names.

5.7.2.10.4 Intraoperative anesthesia administered

Intraoperative anesthesia administered will be summarized by type (General or Regional), ATC level 3 terms and standardized medication names.

5.7.2.10.5 Intraoperative opioid dose

The number and percentage of subjects with intraoperative opioids administered will be summarized by ATC level 3 terms and standardized medication names. The amount of intraoperative opioids administrated (mcgr) will also be summarized.

5.7.2.11 Change in functional status using the ActiGraph wearable device providing measurements for physical activity/function and sleep

The analysis will be performed by an outside vendor and is out of the scope of this SAP.

5.7.2.12 The effects of GRF6021 on plasma proteomics

This analysis will be performed by the Stanford group and by the Sponsor and is out of the scope of this SAP.

5.8 Safety Analyses

Safety analysis include incidence and severity of treatment emergent adverse events, laboratory parameters, vital signs, 12-lead electrocardiogram (ECG), physical examination, pregnancy tests, and prior and concomitant medications. Except where indicated, safety parameters will be summarized by treatment group using the safety population.

5.8.1 Extent of Exposure

The total number of doses received, the actual total volume administrated at each infusion, and the number and percentage of subjects whose infusion stopped or rate changed along with the reason for infusion stop or rate change at each infusion will be summarized.

5.8.2 Adverse Events

Treatment emergent adverse events (TEAEs) will be reported. A treatment emergent adverse event is defined as an AE that occurs on or after the date of the first study treatment.

A summary of all TEAEs by the number and percentages of subjects who experienced any of the following will be provided:

- Any TEAE, including severity (Mild, Moderate, and Severe), relationship to study treatment (Unrelated, Possibly Related, Definitively Related), AEs leading to discontinuation of study participation.
- Any serious TEAE, including severity (Mild, Moderate, and Severe), relationship to study treatment (Unrelated, Possibly Related, Definitively Related), AEs leading to discontinuation of study participation
- Any fatal TEAEs, including relationship to study treatment (Unrelated, Possibly Related, Definitively Related)
- Any event of Special Interest (AESI)

5.8.2.1 Summary for All TEAEs

The number and percentage of subjects who experience TEAEs will be tabulated by system organ class (SOC) and preferred terms (PT) using MedDRA dictionary. Adverse events will be counted only once for a subject within each PT and SOC; thus, since a subject may have more than one PT within a SOC, percentages of PT may not sum to the percentages in the SOC.

For all TEAEs the following will be summarized:

- Overall summary of any TEAEs presented by SOC/PT
- Overall summary of any TEAEs by worst severity presented by SOC/PT
- Overall summary of any TEAEs by closest relationship presented by SOC/PT

A Listing of all AEs will be provided. This listing will include a flag identifying if an AE is a TEAE.

5.8.2.2 Summary for serious TEAEs

For all serious TEAEs the following will be summarized:

- Overall summary of any serious TEAEs by worst severity presented by SOC/PT
- Overall summary of any serious TEAEs by closest relationship presented by SOC/PT

A Listing of all serious AEs will be provided. This listing will include a flag identifying if an AE is a TEAE.

5.8.2.3 Summary for AESI

The following will be considered AEs of special interest (AESI):

- Clinically-significant peripheral edema or pulmonary edema.
- Reduced kidney function (eGFR < 45 mL/min/1.73 m2).
- Suspected transmission of blood-borne infectious agents.

Overall summary of any AESIs by worst severity presented by SOC/PT will be provided.

5.8.2.4 Deaths

Deaths, if any, will be provided in a listing only.

5.8.3 Laboratory Data

Biological samples (e.g. whole blood, serum/plasma, urine) will be collected at Visit 1 Screening and Visit 5 for laboratory evaluations.

Analyses of chemistry, hematology, urinalysis and coagulation tests may consist of the followings:

- Observed values at baseline and Visit 5, along with change from baseline at Visit 5
- Shift tables by analyte and by out of range flag at Visit 5 compared to baseline per subject
- Number and percentage of subjects with abnormal laboratory tests at Visit 1 Screening and Visit 5

The tests will be presented in the original units collected. If the original value contains the "<" sign, AVAL in the ADLB dataset will be derived by extracting the numeric value from LBSTRESC and subtracting 0.1(N+1) to that numeric value, where N is the number of decimals in the numeric value. For example, AVAL will be set to 0.29 if LBSTRESC = "< 0.3". If the LBSTRESC value contains the ">" sign, AVAL in the ADLB dataset will be derived by extracting the numeric value from LBSTRESC and adding 0.1(N+1) to that numeric value, where N is the number of decimals in the numeric value. For example, if LBSTRESC =">1000", AVAL will be set to "1000.1". Subject listings will present all data collected along with toxicity grading.

Results for chemistry, hematology, urinalysis, coagulation, urine drug screening and infectious serology tests will be provided in separate listings. A listing of abnormal laboratory tests will also be provided.

Labs that can indicate serious condition if outside reference range:

Albumin (g/L) < 1
AST > 3xULN
ALT > 3xULN
Alkaline Phosphatase (U/L) > 400
Amylase (U/L) > 120
APTT (sec) > 45
Bicarbonate (mmol/L) < 18 or > 34
Total Bilirubin (mg/dL) > 1.2
BUN (mg/dL) > 45
Ionized Calcium (mg/dL) < 4.5 or > 5.9
Chloride (mmol/L) < 90 or > 110
Creatinine (mg/dL) > 1.4

Creatinine kinase (U/L) > 300
eGFR (mL/min/SA) < 55
GGT (U/L) > 125
Fasting Glucose < Low Range or > High Range
Hematocrit (%) < 30 or > 58
Hemoglobin (g/dL) < 10 or > 20
Hemoglobin A1c (%) > 7.5
Immunoglobulin A < Low Range
Lactate Dehydrogenase (U/L) > 250
Lipase (U/L) > 120
Lymphocytes (%) < 10 or > 60
Magnesium (mg/dL) < 1 or > 2.8
Neutrophils (%) < 20 or > 90
Phosphate (mg/dL) < 1 or > 7
Platelets (x10E3/uL) < 100 or > 800
Potassium (mmol/L) < 3.2 or > 5.6
PT INR > 1.3
PT (sec) > 20
Sodium (mmol/L) < 130 or > 150
Total Lymphs (%) < 8 or > 60
WBC count (x10E3/uL) < 2.5

The mean, median, minimum, maximum, and standard deviation may also be plotted at baseline and Visit 5 by analyte.

5.8.4 Vital Signs

Vital signs, including blood pressure (mmHg), heart rate (beats/min), respiratory rate (breaths/min), and body temperature (C), will be summarized. Vital signs will be collected at Visit 1 Screening and Visit 3 through Visit 5. Observed values will be tabulated at baseline, and at each scheduled post-baseline time point. Change from baseline at each scheduled post-baseline time point will also be summarized. The mean, median, minimum, maximum, and standard deviation may also be plotted at baseline and each scheduled post-baseline time point by parameter.

The vital sign measurements that meet the criteria for blood pressures of special interest (listed below) will also be identified from the vital signs source data. The number and percentage of subjects will be summarized separately by treatment group at each scheduled time point for the following:

- Systolic blood pressure >180 mmHg
- Systolic blood pressure > 200 mmHg
- Systolic blood pressure < 90 mmHg
- Systolic blood pressure > 160 mmHg or < 100 mmHg and > 30% change from baseline
- Diastolic blood pressure > 110 mmHg

- Diastolic blood pressure > 120 mmHg
- Diastolic blood pressure < 50 mmHg
- Diastolic blood pressure > 100 mmHg or < 50 mmHg and > 30% change from baseline

A listing of subjects with any blood pressures of special interest will also be provided. If any SBP or DBP meet the criteria, then all values for both SBP and DBP, regardless of seated or supine, should be listed.

5.8.5 12-lead ECG

A 12-lead ECG will be performed to obtain Heart rate (beats/min), QT interval (msec), and QT interval corrected by the Fridericia formula (QTcF). ECGs will be collected at Baseline Visits 3 and Visit 5. Observed values, along with the change from baseline at Visit 5 will be summarized.

A listing of abnormal 12-lead ECG will also be provided. The mean, median, minimum, maximum, and standard deviation may also be plotted at baseline and Visit 5 by parameter.

5.8.6 Targeted Physical Examination

The body systems examined by the targeted physical examination include lungs/chest, heart, abdomen, and neuro/spine. The number and percentage of subjects with abnormal physical examination results will be summarized by the body systems.

5.8.7 Pregnancy tests

Pregnancy test results will be provided for women of childbearing potential in a listing only.

5.8.8 Prior and Concomitant Medications

The number and percentage of subjects taking any prior and concomitant medications will be summarized by ATC level 3 terms and preferred medication names.

6 Document Version Control

Revision History:

REVISION	RELEASE DATE	AUTHOR	SUMMARY OF CHANGES

Reference

1. Johanna S. Paddison, Tarik Sammour, Arman Kahokehr, Kamran Zargar-Shoshtari, Andrew G. Hill, Development and Validation of the Surgical Recovery Scale (SRS), Journal of Surgical Research, Volume 167, Issue 2, 2011, Pages e85-e91, ISSN 0022-4804, https://doi.org/10.1016/j.jss.2010.12.043. (http://www.sciencedirect.com/science/article/pii/S0022480410019062)
2. Palihnich K, Inouye SK, Marcantonio ER. The 3D CAM Training Manual for Research. 2014; Boston: Hospital Elder Life Program <www.hospitalelderlifeprogram.org>

Appendix A - Programming Specifications for Tables and Listings

The following specifications will be used in the production of tables and listings.

1. Page Setup
   Unless otherwise noted, tables and listings will use landscape orientation. Margins will be at least 3/4 of an inch on the left side of page, at least 3/4 of an inch at the top, and 3/8 of an inch on the other sides.
   Upper left: Sponsor name and protocol number
   - Center: CONFIDENTIAL; Database Download Date: ddmmmyy
   - Upper right: Page number shown as Page n of N. Page numbers should be sequential within a table or listing.

   The footer should include:
   - Left: the name of the SAS program used to generate the output
   - Center: run date/time and the words "by CTDS".
   - Right: output file name.

2. Footnotes
   Unless otherwise specified, footnotes should appear on all pages within the table.

3. Font
   Font will be 9-point Arial, or smaller if needed for space constraints. If possible, small tables should appear on one page. If tables continue on to multiple pages, there should be a page break after an assessment so that all the statistics for an assessment appear on the same page.

4. Tables
   Table titles should reflect the content of the table. Under the main title, in parentheses, the name of the analysis population being summarized should appear.

   4.1 Summary Statistics - Continuous Data

   Unless otherwise noted, the mean and median and confidence interval (CI) of a set of values should be printed out to one decimal place more than the original value. The standard deviation and standard error should be printed out to 2 decimal places more than the original value. The number of subjects on whom the parameter is assessed should appear. Minimum and maximum should be consistent with the original value.

   4.2 Summary Statistics - Categorical Data
   Numbers of subjects are reported as whole numbers. Null counts are represented as 0. Table percentages should be reported to one decimal unless otherwise noted. Null percentages should be reported as 0.0. For all categories, the total number of subjects with data will be presented as N.

5. Subjects Included in Listings
   In general, subject data listings should include all subjects who are randomized. The population flag (Safety/Evaluable/PP) should be included in all listings as a column to indicate which population(s) a subject belongs to. If a listing includes a subset of subjects who meet a certain condition (eg, subjects with SAEs) then this should be clear from the title of the listing. If there is no record for a listing, then a statement, such as There is no serious adverse events in any of the treatment groups, will be presented.