Statistical Analysis Plan

Protocol No.: JS-OAP2-US01, Version 6.0

A Phase 2, Double-Blind, Randomized, Controlled Study to Evaluate the Efficacy and Safety of JointStem, Autologous Adipose Tissue Derived Mesenchymal Stem Cells, in Treatment of Osteoarthritis

Sponsor: Nature Cell Co., Ltd

Protocol Version: JS-OAP2-US01, Version 6.0**,** April 07, 2017

SAP Version No & Date: SAP_ JS-OAP2-US01, Version 2.0, April 05, 2018

Signature Page

***Prepared by*:**

Ami Yu *STAT Manager, HERINGS*	*Date (DD/MM/YYYY)*

Reviewed by:

Byung-Ho Nam *STAT Director, HERINGS*	*Date (DD/MM/YYYY)*
Hugh Lee *Project Director, KCRN Research*	*Date (DD/MM/YYYY)*

Approval by:

Title, Signature *Sponsor*	*Date (DD/MM/YYYY)*

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviation	Term
AdMSC	Adipose Tissue Derived Mesenchymal Stem Cell
AEs	Adverse Events
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
BP	Blood Pressure
BUN	Blood Urea Nitrogen
DSMB	Data and Safety Monitoring Board
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
IKDC	International Knee Documentation Committee
ITT	Intention-to-Treat
KOOS	Knee Injury & Osteoarthritis Outcome Score
LS Means	Least square adjusted means
LOCF	Last-Observation-Carry-Forward
MedDRA	Medical Dictionary For Regulatory Activities
MRI	Magnetic Resonance Imaging
PP	Per Protocol
PT	Preferred Term
PT-INR	Prothrombin Time - International Normalized Ratio
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SI	Système International
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
VAS	Visual Analog Scale
WHO	World Health Organization
WOMAC	Western Ontario and McMaster Universities Arthritis Index

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for Protocol JS-OAP2-US01. This SAP has been developed based on the study protocol (Version 5.0 dated on February 10, 2017) and electronic case report forms (eCRFs) and has been updated according to the amendment of study protocol (Version 6.0, June 26, 2017).

1.1 STUDY OBJECTIVES

1.1.1 Primary Objective

The primary objective of the study is to evaluate the efficacy and safety of autologous adipose tissue derived mesenchymal stem cells (AdMSC) for the treatment of osteoarthritis.

1.2 STUDY ENDPOINTS

1.2.1 Primary Efficacy Endpoints

The co-primary endpoints are:

Using the following three co-primary efficacy variables, the joint functional improvement of knee joint will be evaluated and gatekeeping procedures will be used for these three hypothesis tests.

1) Change of WOMAC score from baseline at Month 6 (JointStem group)

2) Change of VAS score from baseline at Month 6 (JointStem group)

3) MRI improvement evaluation at Month 6 (JointStem group)

   MRI improvement evaluation will be finally analyzed through the following two steps:

   - Radiologist will compare and analyze MRI results before and after the administration of IP and write a report for each subject.
   - Principal Investigator will evaluate MRI results of subjects as one of three conditions (a: improvement, b: no change, c: progress) based on the radiologist's report stated above.

1.2.2 Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

1) JointStem group vs. positive control group
   ① Change of WOMAC score from baseline at Month 6
   ② Change of VAS score from baseline at Month 6
   ③ Change of KOOS from baseline at Month 6
   ④ Change of Lysholm Knee Scoring Scale from baseline at Month 6
   ⑤ Change of IKDC from baseline at Month 6
   ⑥ Change of RAND-36 Score from baseline at Month 6

2) JointStem group
   ① Change of WOMAC score from baseline at Months 9 and 12
   ② Change of VAS score from baseline at Months 9 and 12
   ③ Comparison of MRI improvement evaluations between at Month 6 and at Months 12
   ④ Change of Lysholm Knee Scoring Scale from baseline at Months 6, 9 and 12
   ⑤ Change of KOOS from baseline at Months 6, 9 and 12
   ⑥ Change of IKDC from baseline at Months 6, 9 and 12
   ⑦ Change of RAND-36 Score from baseline at Months 6, 9 and 12

1.2.3 Safety Endpoints

Safety endpoints include:

- Clinical laboratory tests
- 12-lead Electrocardiogram (ECG)
- Physical examinations
- Vital signs
- Adverse events (AEs) and concomitant medications

1.3 SUMMARY OF THE STUDY DESIGN

1.3.1 General Study Design and Plan

This is a double-blind, randomized, controlled study to evaluate the efficacy and safety of JointStem for the treatment of osteoarthritis. Following a 2-week screening period, approximately 30 subjects will be randomly assigned into one of the following two arms in a 2:1 ratio (2 JointStem: 1 positive control):

- JointStem ($1x10^8$ cells/3mL)
- Positive control: Hyaluronic Acid

After each subject completes 6-month visit (Visit 6) and the data management team confirms all individual data have no issue, the individual database will be locked and the blinding will be

open for the statistical analysis. Only subjects who are randomized to JointStem group will then be scheduled for 9-month and 12-month follow-up visits (Visits 7 and 8).

The Schedule of Assessments is presented by Table 1.

Table 1 Schedule of Assessments

	Screen.	Harv.	Treat.	1st Follow-up				2nd Follow-Up	
Blinding	**N/A**	**Double Blinding**						**Open [7]**	
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**ET**	**7**	**8**
Week / Month	**W-7**	**W-5**	**0**	**M1**	**M3**	**M6**	**<M6**	**M9**	**M12 or Final FU[6]**
Informed consent	X								
Medical and Medication History	X	X							
Demographic information	X								
Complete physical examination	X		X		X	X	X[4]	X	X
Vital signs and weight	X	X	X	X	X	X	X	X	X
Hematology, serum chemistry, and urinalysis [1]	X		X	X	X	X	X[4]	X	X
Pregnancy test [2]	X					X	X[4]		X
HIV test for Screening	X								
12-lead ECG	X		X			X	X[4]		X
X-ray	X								
Eligibility confirmation	X	X							
Randomization		X							
Liposuction		X							
Injection of final drug product			X						
AE and concomitant medication assessment			X	X	X	X	X	X	X
MRI scan [3]	X					X	X[5]		X
WOMAC			X	X	X	X	X[4]	X	X
VAS	X		X	X	X	X	X[4]	X	X
Lysholm			X	X	X	X	X[4]	X	X
KOOS			X	X	X	X	X[4]	X	X
IKDC			X	X	X	X	X[4]	X	X
Quality of Life (RAND-36)			X		X	X	X[4]	X	X

1. See Appendix 1 for the complete list of laboratory tests
2. Serum pregnancy test is performed for all females of childbearing potential (please refer to the relevant inclusion criteria)
3. May be performed the day of Visit 2 or after due to scheduling or re-scanning.
4. Performed only if the early termination is after V3.
5. Performed only if the early termination is after V4.
6. Performed if the follow-up is discontinued after V6.
7. Only for JointStem group.

1.3.2 Randomization and Blinding

As this is a double-blinded study, blinding of the drug contents from the subjects, investigators, and other study personnel at each clinical site is necessary. Because two syringes of JointStem and positive control are different, one physician must be only assigned to the role of syringe injection in order to maintain the double-blinding against other site physicians and staffs.

A subject will be assigned with one randomization code on Visit 2 (Week -5). After each subject completes Visit 6 (Month 6) and the data management team confirms all data have no issue, the randomization codes will be open for the statistical analysis.

A code-break envelope for each subject including the randomization code information will be retained by each site and can be opened, revealing the randomization information, for emergency purposes only. Investigator should note that the occurrence of a serious adverse event (SAE) should not routinely precipitate immediate unblinding. An attempt to contact the sponsor must be made prior to unblinding. If unblinding occurs, the subject must be early terminated; a written explanation must be prepared immediately.

1.3.3 Sample Size and Statistical Power Considerations

There was no formal sample size calculation since this is a proof-of-concept study and is not a hypothesis-driven study. The number of enrolled subjects is predefined at n=30, for 20 subjects per study group and 10 subjects per control group.

1.3.4 Interim Analysis

When at least 15 subjects complete 6-month visits (Visit 6, Month 6), one unblinded interim analysis can be performed to evaluate the efficacy and safety of JointStem without the suspension of subject enrollment. No adjustment on the overall alpha-level will be made, as this is not a statistically powered study. The interim analysis is not covered by this SAP.

2. STATISTICAL METHODS

2.1 GENERAL CONSIDERATIONS

In general, continuous variables will be summarized by number of subjects, mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by counts and percentage of subjects in each category. Source data for the summary tables and statistical analyses will be presented as subject data listings, which include data collected on the electronic case report forms (eCRFs) as well as any derived variables for all enrolled subjects.

2.1.1 Reporting Precision

Summary statistics will be presented to the following degree of precision, unless otherwise specified:

Statistics	**Degree of Precision**
Mean, Geometric mean, Median, Quartiles, Confidence limit boundaries	One decimal place more than the raw data.
Standard deviation, Standard error	Two decimal places more than the raw data.
Minimum, Maximum	The same as the raw data.
p-value	Rounded to 4 decimal places and therefore presented as 0.xxxx; p-values smaller than 0.0001 as '<0.0001'; p-values greater than 0.9999 as '>0.9999'.
Percentage	One decimal place. A percentage of 100% will be reported as 100%. Percentages of zero will be reported as 0.

For weight and height, one decimal place will be used for summary statistics. Fractional numeric values will be presented with a zero to the left of the decimal point (for example, 0.12 – 0.30).

2.2 DEFINITIONS OF ANALYSIS POPULATIONS (ANALYSIS SETS)

2.2.1 Intention-To-Treat Population (ITT)

An intention-to-treat (ITT) population will include subjects who received the intra-articular injection and have at least post-injection efficacy measurements.

2.2.2 Per-Protocol Population (PP)

All subjects valid for ITT who completed Visit 6 (Month 6) will be 'valid per protocol' (also called 'valid for efficacy'). Additional criteria may be added prior to unblinding the study database.

2.2.3 Safety Population

A Safety population will include any subject who was randomized and completed Visit 2 (Week -5, at which time adipose tissue was be harvested).

2.3 TIME WINDOWS FOR ANALYSIS

For by visit safety or efficacy summaries, only scheduled visits will be analyzed.

2.4 HANDLING OF MISSING DATA

Last observation carried forward (LOCF) imputation method will be used for efficacy variables, but no adjustments for missing data and no imputation methods are planned for safety variables.

2.5 ANALYSIS SOFTWARE

All summaries and statistical analyses will be generated using SAS® version 9.24 or later.

3. STUDY SUBJECTS

3.1 DISPOSITION OF SUBJECTS

Disposition will be summarized by treatment and total in all enrolled subjects.

The disposition will include the following:

- Subjects Randomized
- Subjects in the ITT Population
- Subjects in the PP Population
- Subjects in the Safety Population
- Subjects completed 6-month follow-up
    - If No, Reasons for withdrawal
- Subjects completed 12-month follow-up
    - If No, Reasons for withdrawal

A listing of dispositions will be provided for all enrolled subjects.

3.2 PROTOCOL DEVIATIONS

The clinical team will identify deviations and the deviations will be recorded into the database. A subject data listing will be provided.

4. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

4.1 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Demographic data (age, gender, race, and ethnicity) and baseline characteristics (including height, weight and body mass index (BMI)) collected at study screening will be summarized using descriptive statistics by treatment group and total for the ITT population.

A subject data listing of demographics and baseline characteristics will be provided.

4.2 MEDICAL HISTORY

Medical history will be coded with the Medical Dictionary for Regulatory Activities (MedDRA) Terminology version 18.1. The medical history conditions will be summarized by treatment group and total for the safety population using frequencies and percentages on the system organ class (SOC) and preferred term (PT). It also will be presented in the data listings.

5. STUDY DRUG AND EXPOSURE

5.1 TREATMENT

A subject data listing will be provided for study drug injection.

5.2 PRIOR AND CONCOMITANT THERAPY

Prior medications are defined as medications that started prior to the intra-articular injection of study drug. Concomitant medications are defined as medications (other than the study drug) taken on or after the intra-articular injection of the study drug during the study. Medications started before the dose of study drug and continuing at the time of the intra-articular injection of study drug are considered both prior medication and concomitant medication.

All investigator terms for medications recorded on the CRF will be coded using the World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) drug dictionary (Version Dec 2015, WHODDE). The number (percentage) of subjects who took prior and concomitant medications will be summarized in the safety population, by ATC Classification, WHO Drug PT, treatment group and total.

A subject data listing of prior and concomitant therapy will be provided.

6. EFFICACY ANALYSES

All efficacy analyses will be performed on the ITT and PP populations.

The primary efficacy hypothesis tests will be performed using a 5% significance level (but, for the MRI improvement evaluation, it shall meet the effective judgment criteria). For the secondary efficacy endpoints, hypothesis tests will be performed individually at the 5% significance level. All hypothesis tests except for the primary efficacy hypothesis tests will be performed with two-sided alternative hypotheses.

6.1 PRIMARY EFFICACY ANALYSIS

For primary efficacy endpoints, the joint functional improvement of knee joint will be evaluated using the following three co-primary efficacy variables and gatekeeping procedures will be used for these three hypothesis tests.

Three hypothesis tests will be performed in the following order: 1) Change of WOMAC score within study group, 2) Change of VAS score within study group and 3) MRI improvement evaluation. If null hypothesis is dismissed at 5% significance level for the first hypothesis test, the second hypothesis will also be tested at 5% significance level. Then, for MRI improvement evaluation, the test will be performed to confirm if the percentage of subjects whose final PI evaluation is 'improvement' meets the effective judgment criteria. If null hypothesis is not dismissed at 5% significance level or it is lower than the effective judgment criteria, the hypothesis test is stopped and it is concluded only previous item of which null hypothesis is dismissed is significant.

1) Change of WOMAC score from baseline at Month 6 (JointStem group)

   Change of WOMAC score from baseline at Month 6 will be tested using paired t-test (one-sided test, α=0.05) with the null hypothesis equal to zero. The change is defined as WOMAC score at Month 6 minus WOMAC score at baseline. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the observed values and change from baseline will be calculated and reported in summary table. In addition, p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of WOMAC score from baseline at Month 6. If change of WOMAC score from baseline at Month 6 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

If null hypothesis is dismissed, the conclusion that the function of knee joint is improved from at least subjects who administer IP can be drawn.

2) Change of VAS score from baseline at Month 6 (JointStem group)

   Change of VAS score from baseline at Month 6 will be tested using paired t-test (one-sided test, α=0.05) with the null hypothesis equal to zero. The change is defined as VAS score at Month 6 minus VAS score at baseline. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the observed values and change from baseline will be calculated and reported in summary table. In addition, p-

value will also be provided. Shapiro-Wilks test will be used to test the normality of change of VAS score from baseline at Month 6. If change of VAS score from baseline at Month 6 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

If null hypothesis is dismissed, the conclusion that the pain of degenerative arthritis is reduced from at least subjects who administer IP can be drawn.

3) MRI improvement evaluation at Month 6 (JointStem group)

MRI improvement evaluation will be finally analyzed through the following two steps:

- Radiologist will compare and analyze MRI results before and after the administration of IP and write a report for each subject.
- Principal Investigator will evaluate MRI results of subjects as one of three conditions (a: improvement, b: no change, c: progress) based on the radiologist's report stated above.

→ The frequency and percentage of subjects whose final PI evaluation is 'improvement' will be calculated per group.

If the percentage of subjects whose final PI evaluation is at least 36%, the conclusion that the damaged cartilage of knee joint is improved in the subjects who administer JointStem can be drawn.

6.2 SECONDARY EFFICACY ANALYSIS

6.2.1 WOMAC Score

WOMAC score will be collected at Visits 3, 4, 5, 6, 7 and 8 (Baseline, M1, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit. Summary statistics of WOMAC score will be presented by visit and treatment group. The change from baseline will also be calculated and represented in summary table.

① Change of WOMAC score from baseline at Month 6 (JointStem group vs. positive control group)
This secondary efficacy endpoint will be analyzed with an ANCOVA with factors of treatment and the baseline of WOMAC score as a covariate. Model-based LS Means of the differences between study and control groups, along with 95% CI and p-value will be provided.

② Change of WOMAC score from baseline at Months 9 and 12 (JointStem group)

Change of WOMAC score from baseline at Months 9 and 12 will be analyzed using paired t-test (two-sided test, α=0.05) with the null hypothesis equal to zero. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the baseline, months 9, months 12 and change from baseline will be calculated and reported in summary table. The p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of WOMAC score from baseline at Months 9 and 12. If change of WOMAC score from baseline at Months 9 and 12 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

6.2.2 VAS Score

VAS score will be collected at Visits 3, 4, 5, 6, 7 and 8 (Baseline, M1, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit. Summary statistics of VAS score will be presented by visit and treatment group. The change from baseline will also be calculated and represented in summary table.

① Change of VAS score from baseline at Month 6 (JointStem group vs. positive control group)
This secondary efficacy endpoint will be analyzed with an ANCOVA with factors of treatment and the baseline of VAS score as a covariate. Model-based LS Means of the differences between study and control groups, along with 95% CI and p-value will be provided.

② Change of VAS score from baseline at Months 9 and 12 (JointStem group)
Change of VAS score from baseline at Months 9 and 12 will be analyzed using paired t-test (two-sided test, α=0.05) with the null hypothesis equal to zero. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the baseline, months 9, months 12 and change from baseline will be calculated and reported in summary table. The p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of VAS score from baseline at Months 9 and 12. If change of VAS score from baseline at Months 9 and 12 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

6.2.3 MRI

① Comparison of MRI improvement evaluations between at Month 6 and at Months 9 and 12 (JointStem group)
MRI improvement evaluation at Month 12 compared to Month 6 will be analyzed using Bowker's test. The frequency and percentage of subjects whose final PI evaluation is

'Improvement' "No change" or "progress" will be calculated at Month 6 and Months 12. The p-value of Bowker's test will be provided.

6.2.4 Lysholm Knee Scoring Scale

Lysholm knee scoring scale will be collected at Visits 3, 4, 5, 6, 7 and 8 (Baseline, M1, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit. Summary statistics of lysholm knee scoring scale will be presented by visit and treatment group. The change from baseline will also be calculated and represented in summary table.

① Change of Lysholm Knee Scoring Scale from baseline at Month 6 (JointStem group vs. positive control group)
This secondary efficacy endpoint will be analyzed with an ANCOVA with factors of treatment and the baseline of Lysholm Knee Scoring Scale as a covariate. Model-based LS Means of the differences between study and control groups, along with 95% CI and p-value will be provided.

② Change of Lysholm Knee Scoring Scale from baseline at Months 6, 9 and 12 (JointStem group)
Change of Lysholm Knee Scoring Scale from baseline at Months 6, 9 and 12 will be analyzed using paired t-test. (two-sided test, α=0.05) with the null hypothesis equal to zero. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the baseline, months 6, months 9, months 12 and change from baseline will be calculated and reported in summary table. The p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of Lysholm Knee Scoring Scale from baseline at Months 6, 9 and 12. If change of Lysholm Knee Scoring Scale from baseline at Months 6, 9 and 12 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

6.2.5 KOOS

KOOS (Symptoms, Pain, ADL, Sport/Rec, QOL) will be collected at Visits 3, 4, 5, 6, 7 and 8 (Baseline, M1, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit. Summary statistics of KOOS will be presented by visit and treatment group. The change from baseline will also be calculated and represented in summary table.

① Change of KOOS from baseline at Month 6 (JointStem group vs. positive control group)
This secondary efficacy endpoint will be analyzed with an ANCOVA with factors of

treatment and the baseline of KOOS as a covariate. Model-based LS Means of the differences between study and control groups, along with 95% CI and p-value will be provided.

② Change of KOOS from baseline at Months 6, 9 and 12 (JointStem group)
Change of KOOS from baseline at Months 6, 9 and 12 will be analyzed using paired t-test (two-sided test, α=0.05) with the null hypothesis equal to zero. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the baseline, months 6, months 9, months 12 and change from baseline will be calculated and reported in summary table. The p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of KOOS from baseline at Months 6, 9 and 12. If change of KOOS from baseline at Months 6, 9 and 12 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

6.2.6 IKDC

IKDC will be collected at Visits 3, 4, 5, 6, 7 and 8 (Baseline, M1, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit. Summary statistics of IKDC will be presented by visit and treatment group. The change from baseline will also be calculated and represented in summary table.

① Change of IKDC from baseline at Month 6 (JointStem group vs. positive control group)
This secondary efficacy endpoint will be analyzed with an ANCOVA with factors of treatment and the baseline of IKDC as a covariate. Model-based LS Means of the differences between study and control groups, along with 95% CI and p-value will be provided.

② Change of IKDC from baseline at Months 6, 9 and 12 (JointStem group)
Change of IKDC from baseline at Months 6, 9 and 12 will be analyzed using paired t-test (two-sided test, α=0.05) with the null hypothesis equal to zero. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the baseline, months 6, months 9, months 12 and change from baseline will be calculated and reported in summary table. The 95% CI and p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of IKDC from baseline at Months 6, 9 and 12. If change of IKDC from baseline at Months 6, 9 and 12 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

6.2.7 RAND-36 Score

RAND-36 Score (Physical Functioning, Role Limitation (Physical), Role Limitation (Emotional), energy/Fatigue, Emotional well-being, Social Functioning, Pain, General Health) at Visits 3, 5, 6, 7 and 8 (Baseline, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit. Summary statistics of RAND-36 Score will be presented by visit and treatment group. The change from baseline will also be calculated and represented in summary table.

① Change of RAND-36 Score from baseline at Month 6 (JointStem group vs. positive control group)
This secondary efficacy endpoint will be analyzed with an ANCOVA with factors of treatment and the baseline of RAND-36 Score as a covariate. Model-based LS Means of the differences between study and control groups, along with 95% CI and p-value will be provided.

② Change of RAND-36 Score from baseline at Months 6, 9 and 12 (JointStem group)
Change of RAND-36 Score from baseline at Months 6, 9 and 12 will be analyzed using paired t-test (two-sided test, α=0.05) with the null hypothesis equal to zero. The number of patients, mean, standard deviation, minimum (min) and maximum (max) values for the baseline, months 6, months 9, months 12 and change from baseline will be calculated and reported in summary table. The 95% CI and p-value will also be provided. Shapiro-Wilks test will be used to test the normality of change of RAND-36 Score from baseline at Months 6, 9 and 12. If change of RAND-36 Score from baseline at Months 6, 9 and 12 doesn't satisfy normal distribution, Wilcoxon signed rank test will be performed.

7. SAFETY ANALYSIS

All safety analyses will be performed for safety population. Statistical tests are not planned for safety variables. Safety assessments include:

- AEs
- Clinical laboratory tests
- ECG
- Physical examinations
- Vital signs

7.1 ADVERSE EVENTS

AEs will be coded by SOC and PT using the MedDRA (Version 18.1). The verbatim term will be included in the AE listing.

A treatment-emergent adverse event (TEAE) is defined as an AE that occur or worsen on or after the injection of study drug and before the end of the study. Only TEAEs will be included in summary tables. All AEs, treatment emergent or otherwise, will be presented in subject data listings.

The TEAEs will be summarized as the number (percentage) of subjects with TEAEs within SOC and PT. Subjects who report the same PT on multiple occasions will be counted once for the PT: under the highest severity (severe > moderate > mild) when summarized by severity and under the closest relationship to study drug when summarized by relationship. If a subject reports multiple PT for a SOC, the subject will be counted only once for that SOC.

TEAEs will be summarized as below.

- An overview table, including number of subjects with
  - TEAEs
  - serious AEs (SAEs)
  - study drug related TEAEs
  - TEAEs by severity
  - TEAEs leading to study discontinuation
  - TEAEs leading to death
- TEAE by SOC and PT
- TEAE by SOC, PT, and Severity
- Study drug related TEAEs by SOC, PT
- SAEs by SOC and PT
- TEAEs leading to study discontinuation by SOC and PT

All AE tables will be sorted by SOC and PT in decreasing frequency of the number and percentage of subjects summarized by treatment group and total.

Missing date will be imputed followed table describes how missing date information will be handled:

Partial Missing Start or Stop Date	Imputed Start Date	Imputed Stop Date
Missing month and day, and the year is present	January 1 of that year or first dose date if the year is the same as the year of first dose date	December 31 of that year
Missing day, but year and month are present	First day of that month or first dose date if the year and month are the same as the year and month of first dose date	Last day of that month
Missing month, but year and day are present	Missing month imputed as January	Missing month imputed as December

7.1.1 Deaths, Serious and Other Significant Adverse Events

The listings of serious AEs, AE leading to study discontinuation, and subjects who died during the study will be provided.

7.2 CLINICAL LABORATORY PARAMETERS

Laboratory assessments (hematology, blood chemistry and urinalysis) are collected at Visits 1, 3, 4, 5, 6, 7 and 8 (Screening, Baseline, M1, M3, M6, M9 and M12) or Early Termination (only if the subject is early terminated after Visit 3) or Final Follow-up Visit:

- Chemistry: Total Bilirubin, Alkaline Phosphatase, Aspartate Aminotransferase (AST) (SGPT), Alanine Aminotransferase (ALT) (SGOT), Blood Urea Nitrogen (BUN), Creatinine, Glucose, Albumin, Total protein, Sodium, Potassium, Chloride, Bicarbonate, Calcium, HCG, HIV.
- Hematology: Hemoglobin, Hematocrit, RBC, WBC, MCV, MCH, MCHC, Neutrophils (absolute), Lymphocytes (absolute), Monocytes (absolute), Eosinophils (absolute), Basophils (absolute), Platelets, PT-INR.
- Urinalysis: Specific gravity, Color, Appearance, pH, Protein, Glucose, Ketones, Bilirubin, Blood, Leukocyte esterase, WBC, RBC, Epithelial cell, Bacteria, Hyaline casts.

All laboratory parameters will be presented in Système International (SI) units. Quantitative results (including actual value and change from baseline) will be summarized using descriptive statistics by baseline and post-baseline visit for each laboratory test group above. Laboratory test results were evaluated as normal, abnormal (not clinically significant) or abnormal (clinically significant). The number and percentage of subjects of each category will be summarized by time point, treatment and total for each parameter.

All laboratory data will be included in the listings. A pregnancy listing will be provided separately.

7.3 VITAL SIGNS, PHYSICAL EXAMINATIONS FINDINGS, ECG

7.3.1 Vital Signs

Vital signs include sitting blood pressure (BP), pulse, temperature, breathing rate and weight. All vital signs will be presented in Système International (SI) units. Quantitative results (including actual value and change from baseline to each post-baseline visit) will be summarized using descriptive statistics by time point, treatment and total for each parameter.

Listings of vital signs with abnormal flags will be provided.

7.3.2 Physical Examinations

Physical examination data will be summarized separately using descriptive statistics by time point, treatment and total, for each component. Listings for physical will be provided.

7.3.3 ECG

The 12-lead ECG parameters including value and changes from baseline will be summarized using descriptive statistics by time point, treatment and total.

Subject listing will be provided. SI units will be used for both summarization and data listings. The critically significant abnormalities will also be flagged in data listings.

8. INTERIM ANALYSES AND DATA AND SAFETY MONITORING BOARD (DSMB)

8.1 INTERIM ANALYSES

When at least 15 subjects complete 6-month visits (Visit 6, Month 6), one unblinded interim analysis can be performed to evaluate the efficacy and safety of JointStem without the suspension of subject enrollment. No adjustment on the overall alpha-level will be made, as this is not a statistically powered study. The interim analysis is not covered by this SAP.

8.2 DSMB

No DSMB is planned.

9. SUMMARY OF MAJOR CHANGES IN THE PLANNED ANALYSES

There is no major change in the planned analyses based on the study protocol (Version 6.0, Apr 7, 2017).

DOCUMENT REVISON HISTORY

DATE	VERSION	AUTHOR/ UPDATED BY	COMMENTS
March 15, 2017	1.0	Hugh Lee	Initial version finalized
April 5, 2018	2.0	Byung-Ho Nam	Updated per protocol v6.0