Statistical Analysis Plan
APSS-44-00

Protocol Title:
A Multicenter, Double-Blind, Randomized, Saline-Controlled Study of a Single, Intra-Articular Injection of Autologous Protein Solution in Patients with Knee Osteoarthritis

Zimmer Biomet, Inc.

Version 7.0

November 6, 2019

Table of Contents

1. Study Objectives

The nSTRIDE Autologous Protein Solution (APS) Kit is designed to be used for the safe and rapid preparation of autologous protein solution (APS) from a small sample of blood at the patient's point of care. The APS is to be injected intra-articularly for the treatment of knee pain associated with osteoarthritis. This study will evaluate the effectiveness and safety of use in a knee OA population.

The WOMAC LK 3.1 is a questionnaire designed to assess osteoarthritic status and progression via patient self-reported pain, function, and stiffness. Thus, the WOMAC is comprised of three subscales. Of the three subscales comprising the WOMAC, the pain subscale score is the basis for the primary endpoint in this study. The primary objective of this study is to determine whether nSTRIDE APS is superior to a saline with regard to the mean improvement from baseline to 12 months in Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) LK 3.1 pain score.

The primary hypothesis to be tested is that the mean improvement in WOMAC pain subscale score (baseline to 12 months) in the APS group will be greater than that of the saline group. Further detail on the primary hypothesis is given in **Section 4.2**.

The impact of the usage of APS on function will be evaluated in support of the primary analysis; however, no formal statistical test for the superiority of function will be incorporated into the study success criteria. Instead, a qualitative assessment of the changes in WOMAC Function over time for APS and Saline will be discussed in order to ensure that the impact of the treatment on function is neutral or positive. Further detail on this qualitative assessment is given in **Section 4.2.**

Secondary objectives of this study include determining whether nSTRIDE APS is superior to saline in improving the WOMAC function and pain subscale scores (as evaluated using the percentage of subjects showing at least the minimal clinically important difference (MCID) and as evaluated using the improvement in mean WOMAC Function subscores at 12 months), OMERACT-OARSI responder rates, analyzing WOMAC pain and function in only the KL-II Subgroup, evaluating superiority of APS over saline in improving Visual Analog Scale (VAS) pain, assessment of the changes in WOMAC Pain scores over time, and evaluation of the usage of restricted and rescue medication. These will be tested in a fixed sequential order pending the rejection of the primary null hypothesis.

Exploratory objectives of this study include determining whether APS is superior to saline with regards to improvement in mean EQ-5D outcomes (12 months minus baseline) and repeated measures of WOMAC Function, WOMAC stiffness, and total WOMAC score. In addition, changes in joint

morphology (determined with MRI images and X-ray images) and laboratory results will be evaluated. Analyses of exploratory endpoints will be done using alpha=0.05 with no adjustment for multiple comparisons and are described in **Sections 4.3** – **4.5**. Subgroup analyses will be conducted to investigate whether the treatment effect varies depending on subjects' initial pain and baseline information.

Finally, the safety profile of nSTRIDE APS will be compared to saline by comparing adverse events and rates between treatment groups.

Testing described in this Statistical Analysis Plan will be performed using SAS 9.2 or a later SAS version.

1.1 Background of the Study

Treatment modalities presently available and directed at treating early to moderate osteoarthritis are palliative, without exception[1]. The nSTRIDE APS Kit builds upon *in vitro* studies, animal studies, and a limited number of human studies designed to understand the causes of osteoarthritis. These research efforts suggest that osteoarthritis is associated with an imbalance in cytokines and growth factors[2]. This imbalance adversely affects cartilage, bone, and soft tissues.

Remodeling of tissues is an ongoing activity in the body. It involves a cycle of tissue breakdown and rebuilding. In osteoarthritis, findings suggest that cytokines associated with cartilage breakdown (Interleukin-1β (IL-1β) and Tumor Necrosis Factor alpha (TNFα)) are too numerous [3-5]. The nSTRIDE APS Kit concentrates beneficial cytokines and growth factors present in the patients' own blood in a way designed to block the activity of these cytokines. In particular, APS concentrates Interleukin-1 receptor antagonist (IL-1ra), soluble Interleukin-1 Receptor II(sIL-1RII), and soluble Tumor Necrosis Factor Receptors I and II(sTNF-RI and sTNF-RII)[6]. This combination of concentrated beneficial factors is intended to act antagonistically to the pro-inflammatory factors causing cartilage degradation and inflammation of the joint (IL-1β and TNFα)[7, 8]. This blockade of inflammation by APS has been demonstrated *in vitro* [9, 10]. Consequently, APS may preserve cartilage and/or improve cartilage health. Imaging analyses comparing a baseline MRI to a 12 month post-injection MRI are aimed at determining whether APS produces structural changes in the joint that can be visualized with this imaging modality. Comparisons between the APS-treated group and the saline-treated group, aimed at determining whether the APS may have any joint preservation potential, are also planned.

1.2 Study Design

The study is a randomized, double-blind study with a planned enrollment of 332 subjects assigned to treatment and control on a 1:1 basis. It is designed to determine whether APS provides an efficacious treatment for knee pain associated with osteoarthritis.

2. Analysis of Baseline Data

Analyses will be done to determine whether randomization succeeded in creating groups that were balanced with regards to key baseline characteristics. Planned tests aimed at determining this are shown in **Table 1**. All tests will be performed under a null hypothesis of no difference between treatment groups.

These tests are intended to be used along with the magnitude/variation as a guideline to see whether any significant baseline imbalances have occurred that would need to be taken into consideration in the analysis of outcomes.

Table 1. Randomization Verification

Variable	Test	Expected Outcome
Gender	Fisher's Exact	No Statistically Significant Difference
Age	T-Test	No Statistically Significant Difference
BMI	T-Test	No Statistically Significant Difference
Race	Likelihood Ratio Chi-Square	No Statistically Significant Difference
Baseline (Screening) WOMAC Pain	T-test	No Statistically Significant Difference

A table showing major and minor protocol deviations will be generated comparing frequency of the occurrence of major and minor deviations between the two treatment groups.

Use of restricted medications will be summarized by type and compared between the two treatment groups.

3. Subject Disposition

A subject disposition table will be created showing the number of subjects in each treatment group that completed the study, with the number of 'in window' and 'all' visits indicated. Visits occurring outside of windows are not considered to be protocol deviations. Visit windows are defined in **Table 2**.

Table 2. Visit Windows

Visit	1 Month	3 Month	6 Month	12 Month
In-Window	± 7 Days	± 14 Days	± 14 Days	± 28 Days
Days since Injection for	23-37	77-105	169-197	337-393

In-Window Visits				
Days since Injection for All Visits	1-61	62-137	138-274	275-548

In addition, optional visits are as follows:

- (Optional) Second Injection (within 14 days of 12 Month visit)
- Second Injection Follow-Up (if needed) will occur within 1 month of second injection (± 7 days)

4. Effectiveness Outcomes

4.1 General Methods

The continuous variables included in the secondary and exploratory analyses will be tested for homoscedasticity using an F-test and also graphically evaluated for normality (boxplots, histograms, and/or normal probability plots). If the p-value for the F-test is less than 0.05 or the plots indicate that the data are extremely non-normal, then the T-tests indicated in **Tables 3a, 3b,** and **3c** will be replaced by an appropriate alternative (Satterthwaite T-test where unequal variances are found or Wilcoxon Mann-Whitney U-test (WMW) where evidence of extreme non-normality is found).

Tables 3a, 3b, and **3c** present a comprehensive list of all planned tests, excluding tests of imaging and blood characterization (which will be presented separately).

4.2 Primary Outcome

The primary objective of this study is to determine whether nSTRIDE APS is superior to a saline with regard to the mean improvement from baseline to 12 months in Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) LK 3.1 pain subscale score.

The primary hypothesis will be tested along with the secondary hypotheses listed in Section 4.2 using a conventional fixed-sequence procedure [11, 12], constructed using a pre-specified order of hypotheses. These tests will be performed at the 0.05 level. If the first null hypothesis is rejected, the second test will be performed at the 0.05 level. If the first null hypothesis is not rejected, the second test, and all subsequent tests, will be performed as exploratory analyses. Since the order of the tests is fixed a priori, and the second hypothesis is tested only if the previous hypothesis has been rejected, the principle of closed testing implies that no adjustment to control the family-wise error rate is necessary. The Fixed-Sequence Method is described as an acceptable method for addressing the multiplicity problem in Section IV.C.5 of the FDA draft guidance entitled, “Guidance for Industry – Multiple Endpoints in Clinical Trials”.

If the primary hypothesis is not tested significantly, all subsequent secondary tests listed in Section 4.2 will be performed as exploratory analyses only.

A two-tailed independent sample T-test will be used to test the primary endpoint. The hypothesis will be:

H_0: $\mu_{APS} = \mu_{Control}$

Versus

H_A: $\mu_{APS} \neq \mu_{Control}$

Where:

μ_{APS}= mean change in WOMAC Pain from baseline to 12 months in the APS group, and

$\mu_{control}$ = mean change in WOMAC Pain from baseline to 12 months in the control group.

A finding of nSTRIDE APS superiority on the mean improvement in WOMAC LK 3.1 Pain along with no corresponding deterioration in WOMAC Function will be considered evidence of nSTRIDE APS efficacy, and the device will be considered efficacious for the treatment of knee pain associated with osteoarthritis.

Table 3a. Planned Primary Efficacy Test

Variable	Test	Expected Outcome
Mean WOMAC Pain Δ (12 Month minus Baseline)	T-Test or WMW	APS Superior

The impact of the usage of APS on function will be evaluated as part of the primary endpoint; however, no formal statistical test will be done as a part of the study success criteria. A qualitative assessment of the changes in WOMAC Function over time for APS and Saline will be discussed, in order to ensure that the impact of the treatment on function is neutral or positive. The following descriptive analyses will be performed for this assessment of function:

(1) A graphical examination of the changes in WOMAC Function for APS and Saline over time, including means and standard error bars.
(2) Descriptive statistics for the mean changes in WOMAC Function over time, separately for each treatment group. These will include mean, median, standard deviation, minimum, maximum, and 95% confidence intervals.

For each analysis, it is expected that the changes in WOMAC function will remain neutral or increase over time within the APS group, and also that the mean changes in WOMAC Function over time will be at least nominally better for the APS group than for Saline.

Note that a repeated measures ANOVA will also be used to assess changes in WOMAC Function over time; this analysis will be exploratory in nature and is described in **Section 4.3.**

4.3 Secondary Outcomes

Secondary objectives of this study include determining whether nSTRIDE APS is superior to saline with regard to the endpoints in **Table 3b**. These secondary hypotheses will be tested using a fixed-sequence procedure, constructed using a pre-specified order of hypotheses. The planned secondary endpoints are ranked in order of importance. The order is shown in **Table 3b**. Subsequent to the rejection of the primary null hypothesis, these tests will be performed at the 0.05 level for each consecutive test until one hypothesis is tested not significantly, after which all subsequent tests will not be performed as secondary analyses, but instead will be performed as exploratory analyses only. As before, since the order of the tests is fixed a priori, and each subsequent hypothesis is tested only if the previous hypotheses have been rejected, the principle of closed testing implies that no adjustment to control the family-wise error rate is necessary.

In particular, it is hypothesized that APS may be superior to Saline with regard to not only an improvement in Pain, but also function thus these comprise the first two secondary endpoints tested. The subjects showing a minimum clinically important difference (MCID) in the WOMAC Function subscale will be referred to as “MCID Function responders.” The proportion of subjects showing a MCID in the WOMAC Pain subscale (i.e. “MCID Pain responders”) will be tested as the second-ranked secondary endpoint.

The development of the MCIDs for which a patient is designated as a “function responder” or a “pain responder” is based on the data from the PROGRESS II pilot study. It is advantageous to use this data for the following reasons:

- The MCID is very much dependent on the choice of clinical score and the scale/version of the score and the follow up time point, so the best approach is to use pilot data where these variables are the same as in the current study.
- The Progress II data allows for development of an MCID that is context-specific, as it has the same or similar patient population, time point, baseline characteristics, baseline symptom severity, and treatments as this PROGRESS IV study. These factors are important to consider in establishing an MCID [13].

- The Pilot study collected a transition question in which patients could rate their improvement as “Very Much Improved”, “Improved”, “Minimally Improved”, “No Change”, “Minimally Worse”, “Much Worse”, and “Very Much Worse”. This question can be used as external criteria to define patients who have experienced a meaningful change in their condition[14].
- Other potential anchor questions are also available in the Pilot study; serving as a way to verify the MCID and make sure it the MCIDs based on these different possible anchor questions are consistent.
- The change that a patient rates as “clinically significant” is dependent on the expectations of the patient[15]. This can vary by treatment as well as by time period.

The MCIDs were derived using an anchor-based method with the patient transition question described above. This question is an appropriate anchor as it is easily clinically interpretable, as well as correlated with the WOMAC Function and Pain scores as shown in the table below. In the Pilot study, the mean improvement in WOMAC Function and Pain scores increases with each increasing patient rating:

TRANSITION QUESTION	N	WOMAC ADL Mean Improvement from Baseline to 12 months	WOMAC Pain Mean Improvement from Baseline to 12 months
Very Much Worse	1	-7.0	-1.0
Much Worse	1	4.5	1.5
Minimally Worse	5	3.8	3.2
No Change	6	7.4	4.2
Minimally Improved	9	15.8	5.8
Much Improved	18	25.3	8.4
Very Much Improved	4	37.8	10.9

Thus the anchor is appropriate as described in methodology in Guyatt, et al.[16] and the MCID is calculated as follows:

The AAOS published a guideline in which calculations of minimum clinically important improvement (MCII) were presented [17]. These calculations of MCII were based on patients with knee osteoarthritis whose final outcome of treatment was “good, satisfactory effect with occasional episodes of pain or stiffness.” The final response to treatment anchored by the baseline value was calculated for each patient. The determinations of clinical significance required

patients in the included studies to achieve a change score comparable to that achieved by 75% of patients reporting good outcomes in the population.

In order to implement this method, the PGI-C response that most closely corresponds to a "good" result was determined. The closest category to the AAOS definition was the "much improved" response (assuming that "minimally improved" corresponds to "fair", and "much improved" corresponds to "good"). The 25th percentile of the distribution of WOMAC Pain scores (improvement from baseline) for these subjects corresponds to a score achieved by 75% of the patients reporting a "much improved (= good)" outcome.

In order to verify the MCID, other anchor question alternatives corresponding to a "good" outcome were explored:

- The SF-36 health transition question, using subjects who respond with "somewhat better"
- The symptom to benefit ratio, which takes into account the symptoms described by the AAOS definition "occasional episodes of pain or stiffness". This includes subjects in the "Moderate/None" category.

The change from baseline in WOMAC function and pain scores for subjects within the appropriate category of each of the three measures were used to calculate and verify the MCID. The MCIDs (i.e. 25th percentile of the distribution of WOMAC Function and Pain Improvement scores) calculated using each of these three anchor questions are presented below.

Transition Question	**MCID for Improvement in WOMAC Function from Baseline to 12 Months**	**MCID for Improvement in WOMAC Pain from Baseline to 12 Months**
Based on 25th percentile of the PGI = "Much Improved" subjects	20.0 points	7.0 points
Based on 25th percentile of the SF-36 health transition question	20.5 points	7.0 points

Transition Question	MCID for Improvement in WOMAC Function from Baseline to 12 Months	MCID for Improvement in WOMAC Pain from Baseline to 12 Months
Based on 25th percentile of the symptom to benefit ratio = "Moderate/None"	16.5 points	7.0 points

The 25th percentiles for each of the three measures were remarkably similar to each other.

To further assess the validity of the MCIDs, the percentage of subjects in each transition category who met the MCID for Function and Pain were calculated, and are as follows:

TRANSITION QUESTION	PERCENT OF SUBJECTS MEETING WOMAC Function MCID	PERCENT OF SUBJECTS MEETING WOMAC Pain MCID
Very Much Worse	0%	0%
Much Worse	0%	0%
Minimally Worse	0%	0%
No Change	16.7%	50%
Minimally Improved	44.4%	55.6%
Much Improved	77.8%	83.3%
Very Much Improved	100%	100%

Based on this data, the MCID is an appropriate differentiator of subjects who are Much improved or Very Much Improved.

Therefore, the MCID Function responder criterion is as follows:

MCID Function Responder:

A subject is considered an MCID Function responder if they show an absolute improvement of ≥ 20.0 points in WOMAC Function from baseline to 12 months.

A two-tailed Fisher's Exact will be used to test the first secondary endpoint. The hypotheses will be:

H_0: $p_{APS} = p_{Control}$

Versus

H_A: $p_{APS} \neq p_{Control}$

Where:

p_{APS}= proportion of MCID function responders in the APS group, and

$p_{control}$= proportion of MCID function responders in the Control group

MCID Pain Responder:

A subject is considered an MCID Pain responder if they show an absolute improvement of ≥ 7.0 points in WOMAC Pain from baseline to 12 months.

A two-tailed Fisher's Exact will be used to test the first secondary endpoint. The hypotheses will be:

H_0: $p_{APS} = p_{Control}$

Versus

H_A: $p_{APS} \neq p_{Control}$

Where:

p_{APS} = proportion of MCID pain responders in the APS group, and

$p_{control}$ = proportion of MCID pain responders in the Control group

The next sequential secondary endpoint will be a comparison of the mean change in WOMAC function score change from baseline to 12 months post-injection.

These analyses of WOMAC function and pain were chosen as the first, second, and third secondary endpoints for analysis because increased pain and decreased function are the predominant clinical findings associated with OA, making both pain and functional improvement important aspects in the treatment of OA. It is anticipated that if treatment with APS decreases pain, it may also increase function. Importantly, the study has sufficient power to test for superiority over the Control for improvement in this key secondary endpoint, in addition to the primary endpoint.

Following the test of the change from baseline in WOMAC function, OMERACT-OARSI Responder Criteria will be applied to both treatment groups, categorizing each patient into one of two categories: responder and non-responder. These results will be tested to determine whether a difference

between treatment groups exists. With respect to the OMERACT-OARSI classification, all measurements where a pain measurement is called for will employ VAS pain. For evaluation of function in the OMERACT-OARSI classification, all measurements will be based on WOMAC function scores. For the global assessment variable, the EQ-5D global assessment VAS scale will be used for OMERACT-OARSI classification. For the function score, absolute change must equal 20% of the total possible score to meet the criteria of an absolute change ≥ 20, and 10% of the total possible score to meet the criteria of an absolute change ≥ 10.

Responders will be defined as subjects who achieved a high degree of improvement in pain or in function (improvement of ≥ 50% and absolute change ≥ 20), ***or*** a moderate degree of improvement in 2 of the 3 response domains (pain, function, global assessment) as follows (**Figure 2**):

(1) VAS pain improvement of ≥20% and absolute improvement of ≥10
(2) WOMAC function improvement of ≥20% and absolute improvement of ≥ 10
(3) EQ-5D global assessment improvement of ≥20% and absolute improvement ≥ 10

A subject who has recurring (two or more) documented uses of rescue medication or restricted medication for index knee OA between the 6 and 12 month visits will be classified as a “non-responder” at the 12 month visit.

The OMERACT-OARSI classification will be calculated as long as there is enough information to calculate per the definition.

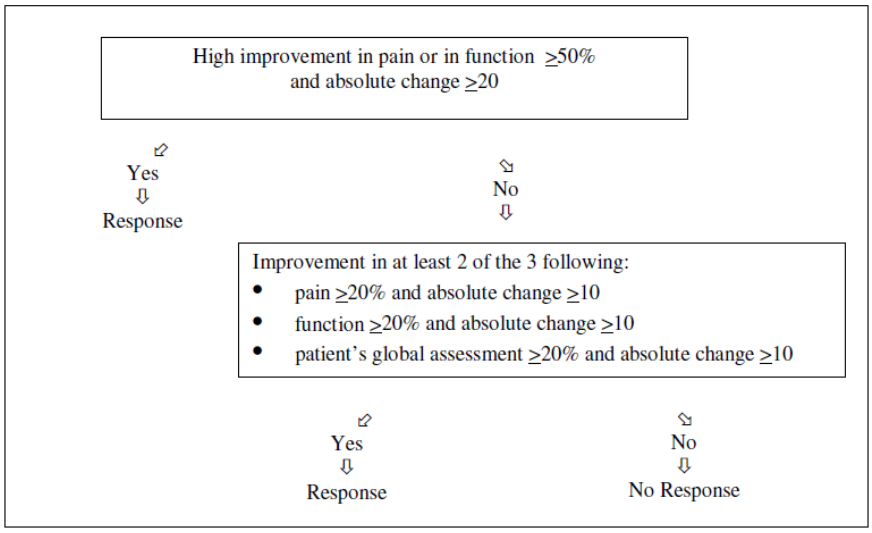

Figure 2. OMERACT-OARSI Responder Criteria [18]

In addition to the test for OMERACT-OARSI classification for the entire study population, the primary efficacy test will be repeated in the subgroup of subjects with Kellgren-Lawrence Grade II (K-L II) OA at baseline. There is supporting evidence in the literature that the lower grade OA patient population may be more responsive to injection therapy [19, 20].

The Kellgren-Lawrence subgroup analyses are ranked below the OMERACT-OARSI classification for the overall study population because the outcomes of the general OA population have greater utility than subgroup analyses.

An analysis of the usage of rescue and restricted medication is considered an important endpoint as it can affect the assessment of pain; thus, this is also a key secondary endpoint. This endpoint is also incorporated as covariate in an ANOVA model to determine if the usage of pain medication has any effect on the mean change in WOMAC Pain from baseline to 12 months.

The analyses of VAS pain from baseline to 12 months are next to last in the planned sequence of secondary analyses, since VAS scores are highly correlated with the WOMAC pain scores[21].

Table 3b. Planned Secondary Efficacy Tests

Order of Testing	Secondary Outcomes	Test	Expected Outcome
1	WOMAC MCID Function Responder rate	Fisher's Exact	APS Superior
2	WOMAC MCID Pain Responder rate	Fisher's Exact	APS Superior
3	Mean WOMAC Function Δ (12 Month minus Baseline)	T-Test or WMW	APS Superior
4	OMERACT-OARSI Responder / Non-responder (12 Month)	Fisher's Exact	APS Superior
5	Mean WOMAC Pain Δ (12 Month minus Baseline) in K-L II Subgroup	T-Test or WMW	APS Superior
6	Mean WOMAC Function Δ (12 Month minus Baseline) in K-L II Subgroup	T-Test or WMW	APS Superior
7	Use of rescue medication (acetaminophen use for index knee OA) over time	Fisher's exact test	APS Superior

Order of Testing	Secondary Outcomes	Test	Expected Outcome
8	Mean WOMAC Pain Δ (12 Month minus Baseline) with Usage of rescue medication within 48 hours of the 12 month visit for knee OA as a covariate	ANOVA	Exploratory
9	Mean WOMAC Pain Δ (12 Month minus Baseline) with Usage of rescue medication within 48 hours of the 12 month visit for any reason as a covariate	ANOVA	Exploratory
10	Mean WOMAC Pain Δ (12 Month minus Baseline) with Usage of restricted medication within 48 hours of the 12 month visit for knee OA as a covariate	ANOVA	Exploratory
11	Mean WOMAC Pain Δ (12 Month minus Baseline) with Usage of restricted medication within 48 hours of the 12 month visit for any reason as a covariate	ANOVA	Exploratory
12	Mean VAS Pain Δ (12 Month minus Baseline)	T-Test or WMW	APS Superior
13	Mean WOMAC Pain changes over time within treatment	Repeated measures ANOVA	Significant change from baseline

4.3.1 Exploratory Outcomes (Questionnaire)

Exploratory outcomes include determining whether APS is superior to saline with regards to improvement in repeated measures of WOMAC Pain and WOMAC Function, WOMAC stiffness, and total WOMAC score. In each test associated with these measures, change from baseline will be evaluated. Change in EQ-5D measures from baseline to 12 months post-procedure will

be evaluated and, separately, in repeated measures at 1, 3, and 6 months post-procedure. A statistically significant difference favoring APS is expected at 12 months post-procedure, and a trend favoring APS is expected in the repeated measures tests covering 1, 3, and 6 months for all tests. Where repeated measures tests are statistically significant, tests at individual time points will be performed. All tests will be conducted at $\alpha = 0.05$ and will not be adjusted for multiple comparisons as these are exploratory outcomes.

Table 3c. Planned Exploratory Efficacy Tests: Questionnaire Data

Exploratory Outcomes	Test	Expected Outcome
Mean WOMAC Pain Δ (Percent change from baseline to12 Month)	T-Test or WMW	APS Superior
Mean WOMAC Function Δ (Percent change from baseline to 12 Month)	T-Test or WMW	APS Superior
Mean WOMAC Function changes over time within treatment	Repeated measures ANOVA	Significant change from baseline
Mean EQ-5DΔ Global Assessment VAS (12 Month minus Baseline)	T-Test or WMW	APS Superior
Mean EQ-5DΔ Global Assessment VAS (Percent change from baseline to12 Month	T-Test or WMW	APS Superior
Mean EQ-5D Global Assessment VAS changes over time within treatment	Repeated measures ANOVA	Significant change from baseline
Mean EQ-5D Δ Single Index Value (12 Month minus Baseline)	T-Test or WMW	APS Superior
Mean EQ-5DΔ Single Index Value (Percent change from baseline to 12 Month	T-Test or WMW	APS Superior
Mean EQ-5D Single Index Value changes over time within treatment	Repeated measures ANOVA	Significant change from baseline

Exploratory Outcomes	Test	Expected Outcome
EQ-5D Dimensions (Mobility, Self-Care, Usual Activities, Pain/Discomfort, Anxiety/Depression) (12 Month)	Likelihood Ratio Chi-Square	APS Superior
Mean WOMAC Pain Δ (1, 3, & 6 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior
Mean WOMAC Function Δ (1, 3, & 6 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior
Mean VAS Pain Δ (1, 3, & 6 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior
MCID WOMAC Function Responder (1, 3, & 6 Months)	Fisher's Exact Family α = 0.05	Trend APS Superior
MCID WOMAC Pain Responder (1, 3, & 6 Months)	Fisher's Exact Family α = 0.05	Trend APS Superior
OMERACT-OARSI Responder / Non-responder (1, 3, & 6 Months) **Traditional published criteria; i.e. not including multiple restricted/ rescue medication usage between 6-12 months	Fisher's Exact Family α = 0.05	Trend APS Superior
Total Mean WOMAC Δ (1, 3, & 6 and 12 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior
Mean WOMAC Stiffness Δ (1, 3, & 6 and 12 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior

Exploratory Outcomes	Test	Expected Outcome
Mean EQ-5DΔ Global Assessment VAS (1, 3, & 6 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior
Mean EQ-5D Δ Single Index Value (1, 3, & 6 Months)	Repeated Measures ANOVA or Friedman Test	Trend APS Superior
EQ-5D Dimensions (Mobility, Self-Care, Usual Activities, Pain/Discomfort, Anxiety/Depression) (1, 3, & 6 Months)	Likelihood Ratio Chi-Square	Trend APS Superior
Mean WOMAC Pain Δ (12 Month minus Baseline) by Treatment and Site	ANOVA	No Site Effect
Mean WOMAC Function Δ (12 Month minus Baseline) by Treatment and Site	ANOVA	No Site Effect
Mean WOMAC Function Δ (12 Month minus Baseline) with Usage of rescue medication within 48 hours of the 12 month visit for knee OA as a covariate	ANOVA	Exploratory
Mean WOMAC Function Δ (12 Month minus Baseline) with Usage of rescue medication within 48 hours of the 12 month visit for any reason as a covariate	ANOVA	Exploratory

Exploratory Outcomes	Test	Expected Outcome
Mean WOMAC Function Δ (12 Month minus Baseline) with Usage of restricted medication within 48 hours of the 12 month visit for knee OA as a covariate	ANOVA	Exploratory
Mean WOMAC Function Δ (12 Month minus Baseline) with Usage of restricted medication within 48 hours of the 12 month visit for any reason as a covariate	ANOVA	Exploratory
WOMAC Pain Δ: Baseline-12M	Summary Statistics and ANCOVA Model*	Exploratory
WOMAC Pain Δ: Baseline-6M	Summary Statistics and ANCOVA Model*	Exploratory
Percent of Subjects in each Treatment/Cross Over Category	Chi-Square	Exploratory

*Analysis of Covariance models will be used in order to determine if changes in WOMAC Pain score from Baseline to12 months (first model) or from 6 to 12 months (second model) differ according to what type of 2nd injection they chose to receive (β_3), and whether this difference varies for APS and HA ($\beta_2 * \beta_3$). The model will be as follows:

$Y=\beta_0+ \beta_1+ \beta_2+ \beta_3+ \beta_2*\beta_3$

Where

β_0= Intercept

β_1= Baseline WOMAC Pain score

β_2 = Treatment Group

β_3= 2nd Injection (None, Same, Crossover)

Y= Change in WOMAC Pain

4.4 Exploratory Outcomes (Imaging)

This study will evaluate changes in joint morphology, determined with MRI images and X-ray images. All imaging results will be evaluated by a central laboratory. MRI and X-rays will be taken prior to the injection procedure and at 12months post-injection.

4.4.1 X-Ray Outcomes

X-rays are evaluated for Kellgren-Lawrence (K-L) grade and absence of severe osteoarthritis by the central laboratory for confirmation of eligibility.

Post-injection X-rays taken at 12 months will receive K-L grades that will be compared to baseline. For each subject, results will be categorized as (1) worse—a higher Kellgren-Lawrence grade at 12 months compared to baseline, (2) no change in grade from baseline to 12 months or (3) improvement—a decrease in grade from baseline to 12 months. An analysis of the percentage of subjects in each treatment group who fall into each of these three categories will be done using a Likelihood Ratio Chi-square test. This test will also be repeated within subgroups defined by the baseline K-L grade.

Measurements of medial and lateral Joint Space Width (JSW) will be performed to assess narrowing over time in the medial and lateral compartments of the treated joint. Joint Space Narrowing (JSN) will be calculated for each compartment as the change in JSW between the 12 Month visit and the Baseline visit.

A list of planned statistical tests on imaging data is presented in **Table 4a**. All tests will be conducted at $\alpha = 0.05$ and will not be adjusted for multiple comparisons as these are exploratory outcomes. In addition to these tests, regression methods will be used to determine whether imaging results are associated with MCID responder status, change in WOMAC pain or Function scores from baseline to 12 months, or with OMERACT-OARSI categorization at 6 and 12 months (see **Table 4c**).

Table 4a. Planned Statistical Tests for X-ray Results

Variable	Test
X-ray	
Kellgren-Lawrence (baseline, 12 months, and change)	Likelihood Ratio Chi-Square for each
Mean JSW at 12 Months (medial and lateral compartments)	T-Test or WMW for each compartment

Variable	Test
X-ray	
Mean JSN (12 months minus baseline) (medial and lateral compartments)	T-Test or WMW for each compartment
Change in K-L by baseline K-L grade	Likelihood Ratio Chi-Square for each subgroup

4.4.2 MRI Outcomes

Imaging will be evaluated in an effort to determine whether APS produces measureable physiological changes in the progression of osteoarthritis. This study will evaluate change in joint morphology by comparing baseline and 12 month high resolution MRI imaging changes between treatment groups (APS knee joint injection versus a saline control knee joint injection).

Quantitative measurements, including Coronal Lesion Size, Sagittal Lesion Size, T2 Relaxation Map, and Normalized T2 Relaxation Time, will be collected at baseline and 12 months. Lesion sizes will be reported in millimeters, and normalized T2 values will be reported as unitless decimals.

The following qualitative, visual assessments will be performed:

- Subchondral Fracture
- Bone Marrow Lesion Size
- Cyst Percentage
- Osteophytes
- Cartilage Loss
- Meniscal Extrusion
- Meniscal Morphology
- Synovial Thickening & Joint Effusion
- Peri-Articular Features
- Loose Bodies
- Additional Observations

A detailed classification system for each qualitative assessment is provided in the Image Review Charter for this study.

A list of planned statistical tests on imaging data is presented in **Table 4b**. All tests will be conducted at $\alpha = 0.05$ and will not be adjusted for multiple comparisons as these are exploratory outcomes. In addition to these tests, regression methods will be used to determine whether imaging results are associated with MCID responder status, change in WOMAC Pain and Function scores from baseline to 12 months, or with OMERACT-OARSI categorization at 6 and 12 months (see **Table 4c**).

Table 4b. Planned Statistical Tests for MRI Results

MRI / MOAKS Scales	Test
Coronal Lesion Size	T-Test or WMW
Δ Coronal Lesion Size	T-Test or WMW
Sagittal Lesion Size	T-Test or WMW
Δ Sagittal Lesion Size	T-Test or WMW
Normalized T2 Relaxation time	T-Test or WMW
Δ Normalized T2 Relaxation time	T-Test or WMW
Subchondral Fracture	Fisher's Exact
Derived* Bone Marrow Lesion Changes	Likelihood Ratio Chi-Square
Derived* Cyst Percentage Changes	Likelihood Ratio Chi-Square
Derived* Osteophytes Changes	Likelihood Ratio Chi-Square
Derived* Cartilage Loss Changes	Likelihood Ratio Chi-Square
Derived* Meniscal Extrusion Changes	Likelihood Ratio Chi-Square
Derived* Meniscal Morphology Changes	Likelihood Ratio Chi-Square
Synovial Thickening &Joint Effusion	Likelihood Ratio Chi-Square
Peri-Articular Features	Fisher's Exact
Loose Bodies	Fisher's Exact

***Change from baseline to 12 Month will be assigned to one of six categories as defined in the Image Review Charter for this study.**

Table 4c. Planned Statistics for Imaging/ Questionnaire Results

Imaging Variables	Dependent (Outcome) Variable	Statistic
All Measured MRI, X-ray variables, and Treatment Assignment	MCID WOMAC Function responder Category at 6 and 12 Months	Logistic Regression
All Measured MRI, X-ray variables, and Treatment Assignment	Mean WOMAC Pain Δ (Baseline to 12 Months)	ANOVA
All Measured MRI, X-ray variables, and	Mean WOMAC Function Δ (Baseline to 12	ANOVA

Imaging Variables	Dependent (Outcome) Variable	Statistic
Treatment Assignment	Months)	
All Measured MRI, X-ray variables, and Treatment Assignment	OMERACT-OARSI Category at 6 and 12 Months	Logistic Regression

4.5 Exploratory Outcomes (Laboratory Tests of Blood and APS)

The nSTRIDE APS Kit concentrates cytokines and growth factors present in the patient's blood that are believed to be beneficial, based on prior evidence. Therefore, the relationships between APS cytokine and growth factor concentrations and osteoarthritis-related questionnaire outcomes will be evaluated with a focus on the APS treatment group.

Laboratory tests of blood and APS will be evaluated by a central laboratory. Specimens will be taken from the injection visit. A separate APS sample, not intended for injection, will be created at the injection visit. This sample will be evaluated for both APS and Saline control subjects; in saline control subjects, both samples of APS will be combined and sent to the central laboratory for evaluation. The tests indicated in **Tables 5a** and **5b** will be performed separately for APS and Saline subjects. As indicated, Fisher's z transformations will be used to determine whether differences between treatment group correlations exist, where appropriate.

The tests in **Table 5c** are aimed at determining whether blood characteristics are related to APS output. Clinical outcomes are not included in these tests, and APS will be created for APS and Saline subjects (although it will not be injected in Saline patients). Accordingly, these tests will be performed on pooled data from APS and Saline subject specimens.

All variables where Pearson-r tests are planned will be assessed for normality using graphical methods. Where the data are severely non-normal, a Spearman-rho test will be used instead of a Pearson-r. All Fisher's z transformations comparing correlations between treatment groups will be two-tailed.

All tests will be conducted at $\alpha = 0.05$ and will not be adjusted for multiple comparisons as these are exploratory outcomes.

Table 5a. Planned Statistical Analyses for APS / Blood Results (All-inclusive and Anti-inflammatory)

APS / Blood Variable	Outcome Variable(s)	Statistic	Expected Outcome	
APS/Blood	Questionnaire		APS	Saline
All Measured APS Variables	MRI Progression Variables	Polyserial Correlation	Unknown	Unknown
All Measured Blood Variables	MRI Progression Variables	Polyserial Correlation	Unknown	Unknown
All Measured APS Variables	Δ from baseline to all post-injection time points in: WOMAC Pain, Function, Stiffness and Total Score; VAS Pain, EQ-5D Index and EQ-VAS	Canonical Correlation	Unknown	Unknown
All Measured Blood Variables	Δ from baseline to all post-injection time points in: WOMAC Pain, Function, Stiffness and Total Score; VAS Pain, EQ-5D Index and EQ-VAS	Canonical Correlation	Unknown	Unknown
All Measured APS Variables	Responder status: MCID WOMAC Function and OMERACT-OARSI Category at all post-injection time points	Polyserial Correlation	Unknown	Unknown
All Measured Blood Variables	Responder status: MCID WOMAC Function and OMERACT-OARSI Category at all post-injection time points	Polyserial Correlation	Unknown	Unknown
APS IL-1ra/IL-1β Concentration Ratio	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient + Fisher's z	Unknown	No Correlation
Whole Blood IL-1ra/IL-1β Concentration Ratio	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient + Fisher's z	Unknown	No Correlation

APS / Blood Variable	Outcome Variable(s)	Statistic	Expected Outcome	
APS/Blood	Questionnaire		APS	Saline
APS IL-1ra$_{pg/ml}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
Whole Blood IL-1ra$_{pg/ml}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
APS IL-1ra/IL-1β Concentration Ratio	Responder status: MCID WOMAC Function and OMERACT-OARSI Category at all post-injection time points	Point-biserial correlation coefficient	Positive Correlation	No Correlation
APS IL-1ra/IL-1β Concentration Ratio ≥ 1000 (Y/N)	Responder status: MCID WOMAC Function and OMERACT-OARSI Category at all post-injection time points	Point-biserial correlation coefficient	Positive Correlation	No Correlation

Table 5b. Planned Statistical Analyses for APS / Blood Results (WBCs and Growth Factors)

APS / Blood Variable	Outcome Variable(s)	Statistic	Expected Outcome	
APS/Blood	Questionnaire/Imaging		APS	Saline
APS WBC$_{k/µl}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Positive Correlation	No Correlation
Whole Blood WBC$_{k/µl}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Positive Correlation	No Correlation
APS TGF-β1$_{pg/ml}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
Whole Blood TGF-β1$_{pg/ml}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation

APS / Blood Variable	Outcome Variable(s)	Statistic	Expected Outcome	
APS/Blood	Questionnaire/Imaging		APS	Saline
APS IGF-$1_{pg/ml}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
Whole Blood IGF-$1_{pg/ml}$ Concentration	Δ WOMAC Pain and Function from baseline to all post-injection time points	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
APS IGF-$1_{pg/ml}$ Concentration	Joint Space Narrowing (each compartment)	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
Whole Blood IGF-$1_{pg/ml}$ Concentration	Joint Space Narrowing (each compartment)	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
APS TGF-$\beta1_{pg/ml}$ Concentration	Joint Space Narrowing (each compartment)	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
Whole Blood TGF-$\beta1_{pg/ml}$ Concentration	Joint Space Narrowing (each compartment)	Pearson-r correlation coefficient+ Fisher's z	Unknown	No Correlation
APS IGF-$1_{pg/ml}$ Concentration	MRI progression variables	ANOVA	Unknown	No Differences
Whole Blood IGF-$1_{pg/ml}$ Concentration	MRI progression variables	ANOVA	Unknown	No Differences
APS TGF-$\beta1_{pg/ml}$ Concentration	MRI progression variables	ANOVA	Unknown	No Differences
Whole Blood TGF-$\beta1_{pg/ml}$ Concentration	MRI progression variables	ANOVA	Unknown	No Differences

Table 5c. Planned Statistical Analyses Comparing APS / Blood Results

APS / Blood Variable	Outcome Variable(s)	Statistic	Expected Outcome
Blood WBC Concentration to IL-1ra Concentration ratio	APS WBC Concentration to IL-1ra Concentration ratio	Paired-sample t-test	No Difference
Blood Neutrophil Concentration to IL-1ra Concentration ratio	APS Neutrophil Concentration to IL-1ra Concentration ratio	Paired-sample t-test	No Difference
Blood Monocyte Concentration to IL-1ra Concentration ratio	APS Monocyte Concentration to IL-1ra Concentration ratio	Paired-sample t-test	No Difference

The ratio of IL-1ra to WBC, neutrophil, and monocyte concentration in whole blood will be compared to the same ratio in APS to determine whether a difference exists in these ratios between whole blood and APS, which would indicate a difference in the amount of IL-1ra per cell in the solution. These ratios will be compared using a paired-sample, two-tailed T-test, with an expected outcome of no evidence of difference.

Each APS sample will be analyzed to determine if it meets the following criteria:

- WBC count ≥ 10k/µl
- Deliverable dose of IL-1ra ≥ 6 ng
- IL-1ra:IL-1β concentration ratio ≥ 100
- sIL-1RII:IL-1β concentration ratio ≥ 1
- sTNF-RII:TNFα concentration ratio ≥ 1

If any samples from APS-treated patients do not meet one or more of the above-listed criteria, a listing will be created to show the change from baseline to 12 months in WOMAC pain and function scores and OMERACT-OARSI responder status at 12 months for those subjects with samples not meeting the criteria. Subjects producing APS volume less than 1 ml from either processed device will also be included in the listing. Descriptive statistics will be performed. If a sufficient number of samples do not meet one or more of the criteria, analyses will be performed to determine if a relationship exists between these criteria and the key study outcomes. Specifically:

- The association between ability to meet these criteria and the change in WOMAC pain and (separately) function scores from baseline to 12

months will be analyzed using an ANOVA model. The dependent variable will be the change in WOMAC pain or function score, and the independent variable will be whether or not the criteria were met (Yes/No).

- The association between the ability to meet these criteria and the MCID WOMAC Function Responder status, and OMERACT-OARSI responder status at 12 months will be analyzed using logistic regression models. The dependent variable will be the responder status at 12 months, and the independent variable will be whether or not the criteria were met (Yes/No).

4.6 Subgroup Analyses

Clinically relevant differences in treatment effect are not anticipated across age, gender, race, or other subgroups, and the primary analysis will not be stratified by any subgroups. However, subgroup analyses of primary and secondary efficacy and safety endpoints will be performed in an exploratory fashion as specified below. These analyses will be performed for each of the following subgroups:

- Age (treated as continuous)
- Age (< Median vs. ≥ Median)
- Gender (Male/Female)
- Race (White (Hispanic), White (non-Hispanic), African-American, Native American, Asian or Pacific Islander, Other, or Not Specified).
- Baseline KL Grade
- Presence of Bone Marrow Lesions at baseline
- Site
- Baseline WOMAC Pain (< Median vs. ≥ Median)
- Usage of rescue or restricted medications prior to the 12M visit
- Presence of Contralateral Knee Pain prior to the 12M visit
- KL improvement status (improved, worsened, unchanged)

The statistical models in **Table 6** will be performed for each subgroup. If the Treatment*Subgroup interaction is statistically significant ($p < 0.05$) and/or clinically meaningful, further analysis will be performed to determine the particular subgroup(s) in which the treatment effect differs, and the impact of this will be assessed and described. The least squared means of each treatment*subgroup combination will be output from the model.

Table 6. Subgroup Analysis

Dependent Variable(s)	Statistical Test
Mean WOMAC Pain Δ (12 Month minus Baseline)	Analysis of Covariance Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction

Dependent Variable(s)	Statistical Test
WOMAC MCID Function Responder rate	Logistic Regression Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
WOMAC MCID Pain Responder rate	Logistic Regression Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
Mean WOMAC Function Δ (12 Month minus Baseline)	Analysis of Covariance Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
OMERACT-OARSI Responder / Non-responder (12 Month)	Logistic Regression Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
Mean WOMAC Pain Δ (12 Month minus Baseline) in K-L II Subgroup	Analysis of Covariance Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
Mean WOMAC Function Δ (12 Month minus Baseline) in K-L II Subgroup	Analysis of Covariance Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
Mean VAS Pain Δ (12 Month minus Baseline)	Analysis of Covariance Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
Occurrence of one or more SAE (Yes/No)	Logistic Regression Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction
Occurrence of one or more Device-Related AE (Yes/No)	Logistic Regression Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction

Dependent Variable(s)	Statistical Test
Rescue Medication Use for Index Knee OA Pain (Yes/No)	Logistic Regression Independent variables: Treatment, Subgroup, Treatment*Subgroup Interaction

4.7 Analysis of Data after 2nd Injection

Data will be captured approximately 1 month after any second injection that occurs. A table will be generated showing how many patients in each treatment group received a second injection. Questionnaire and adverse event data collected following the second injection will be presented separately.

In addition, the following comparisons will be made:

Table 7. Analysis of Data after 2nd Injection

Dependent Variable(s)	Statistical Test
Percent of Subjects electing to receive 2nd Injection	Chi-Square (APS vs Control)
Percent of Subjects with a Device-Related AE	McNemar Test on Paired Proportions (Before 2nd Injection vs After 2nd Injection)

5. Safety Outcomes

5.1 Adverse Events

Adverse Event definitions and assessment descriptions are included in the protocol (APSS-44-00, Section 4).

Adverse events will be designated into MedDRA categories within the EDC system. Specific adverse events are expected to occur at low frequencies, and therefore no differences between groups are expected. Accordingly, the *a priori* tests planned will determine whether the overall rate of AEs, SAEs, ADEs, and UADEs differs between treatment groups, and whether differences between groups exist with respect to device relatedness and severity. Results for each category will be tabulated by treatment group. In this tabulation the number of subjects with one or more adverse events fitting into an individual adverse event (AE) category will be displayed beside the count of subjects with no AEs fitting into that category. A Fisher's Exact test will be performed to determine whether treatment groups differ. A listing of all AEs will be provided, and a narrative of each serious AE will be generated.

In addition, a time-course distribution of adverse events will be created based on contiguous intervals between planned visits, as shown in **Table 8**. This chronology will include AEs, SAEs, ADEs, and UADEs. The time-course will also be displayed after stratifying events by severity and relatedness to the device.

Table 8. Windows for Time Course Distribution of AEs

Interval	Beginning of Interval (# Days)	End of Interval (# Days)
Injection	0*	15
1 Month	16	61
3 Month	62	137
6 Month	138	274
12 Month	275	2nd Injection or Endpoint
2nd Injection	2nd Injection	15 Days after 2nd Injection
2nd Injection FU	16 Days after 2nd Injection	Endpoint

*Day of injection

5.2 Rescue Medication and Withdrawal for Treatment of Index Knee OA

A time-course tabulation showing rescue medication (acetaminophen) use (for index knee OA) by treatment group will be generated. In addition the percentages of subjects requiring rescue medication for their index knee OA during the course of follow-up (up to 12 months) will be compared for APS vs. Control using a Fisher's Exact test.

Subjects who receive invasive treatment of their index knee and/or continued use of restricted medications during follow-up, as evidenced by multiple protocol deviations, may be withdrawn from the study on the basis of insufficient compliance with the protocol (APSS-44-00, Section 2.7).

These subjects will undergo a blinded review by an independent data monitoring committee to assess whether the withdrawal was related to the index knee OA.

Subjects who are withdrawn prior to completing the 12 month follow-up visit due to reasons related to the index knee OA will be considered "non-responders", and the subjects' baseline scores will be used in the primary analysis of WOMAC Pain, and in the sensitivity analyses described in **Section 6.5**. Missing data from these subjects will not be imputed in any other analyses. Primary endpoint data from subjects who are withdrawn due to reasons unrelated to the index knee OA will be treated as "missing" and imputed as described in **Section 6.4**.

In order to summarize the time to failure ('non-response' designation) for APS and control groups, a Kaplan Meier analysis will be used. A log-rank test will be used to compare the survival curves for the two groups. Subjects who are

withdrawn due to reasons related to the index knee OA will be considered "failures", with failure time equal to the date of withdrawal. Subjects who are withdrawn for reasons unrelated to the index knee OA will be censored at the date of withdrawal. All other subjects will be censored at the date of their last follow-up visit.

6. Study Populations

Eligible patients must have symptomatic osteoarthritis (OA) in one knee (only), and must have been unable to get satisfactory pain relief with prior conservative treatment. Patients of both genders ≥ 21 years and ≤ 80 years of age will be eligible. Patients must have a Kellgren-Lawrence osteoarthritis grade of 2 to 4 and an absence of severe osteoarthritis, confirmed by a central laboratory from pre-procedure X-ray. All inclusion and exclusion criteria are presented in the study protocol.

6.1 Sample Size

The sample size for this study is 332 subjects, to be enrolled and randomized into one of two treatment groups (APS and saline).

This sample size will give sufficient power to test the current primary hypothesis as well as the first two secondary hypotheses, as described below.

The sample sizes needed for the primary and first two secondary endpoints were calculated, and the largest was chosen as the sample size for this study.

Data used to estimate sample size were from Zimmer Biomet study APSS-33-00, a randomized study comparing APS to saline in human subjects. Estimates of expected mean improvements in WOMAC Pain are shown in **Table 9a,** proportions of WOMAC Function MCID responders at 12 months from this study are shown in **Table 9b**, and estimates of expected mean improvements in WOMAC Function are shown in **Table 9c.** Sample sizes were calculated using SAS 9.4.

Table 9a. WOMAC Pain Sample Size Based on Results from APSS-33-00

Variable	Difference in Mean Improvement* Seen in Pilot Study (APS-Saline)	StdDev of Mean Improvement* seen in Pilot Study	Type I Error	Power	Sample Size Per Group (With- 15% Attrition) Total Study Sample Size
WOMAC Pain	2.41	4.01	0.05	0.90	60 (71) 142

***12M-Screening**

Table 9b.WOMAC Function MCID Responder Sample Size Based on Results from APSS-33-00

Variable	Proportion of MCIDSeen in Pilot Study (APS)*	Proportion of MCIDSeen in Pilot Study (Saline)*	Type I Error	Power	Sample Size Per Group (With 15% Attrition) Total Study Sample Size
WOMAC Function	83.3%	57.1%	0.05	0.90	69(82) 164

***12M-Screening**

Table 9c.WOMAC Function Sample Size Based on Results from APSS-33-00

Variable	Difference in Mean Improvement* Seen in Pilot Study (APS-Saline)	StdDev of Mean Improvement* seen in Pilot Study	Type I Error	Power	Sample Size Per Group (With 15% Attrition) Total Study Sample Size
WOMAC Function	6.04	15.58	0.05	0.90	141 (166) 332

***12M-Screening**

The highest required sample size was 141 per group. An increase of 15% for attrition [22] yields a sample size of 166 per group (N=332 total).

Thus, the original sample size of 332 subjects is sufficient to test the current study primary hypothesis, as well as the first two secondary hypotheses. A maximum of 30 sites will participate in the study. Enrollment will be stopped at any site reaching a limit of 30% of planned study enrollment (332 total subjects).

6.2 Treatment Assignment

Subjects will be assigned 1:1 to either APS or a saline control group. The study will be double-blind. Assignment will be stratified by site with random block sizes.

6.3 Analysis Populations Assignment

A list of protocol violations will be evaluated to determine subjects who (1) violated one or more eligibility criteria, and (2) subjects with major protocol deviations that impact patient safety or the scientific validity of the study comparisons, or have the potential to do so. This will be determined by a blinded review of the protocol deviations.

Results will be tabulated. For the primary endpoint, three analyses will be performed: (1) intent-to-treat (ITT), (2), modified intent-to-treat (mITT), and (3) per protocol (PP), as defined below. The ITT analysis will be considered the primary analysis of the study, and the mITT and PP analyses will be considered as sensitivity analyses.

Secondary endpoints will be analyzed using the per protocol group, with no imputation for missing data. Safety evaluations will be performed on the intent-to-treat group, with no imputation for missing data. Further detail regarding these analysis groups is given below.

Intent-to-Treat (ITT)

The ITT population includes all randomized subjects, regardless of whether they received an injection or not. This will be the population used for any data listings, for the primary study analysis, for safety analyses, and for demographic data summarized according to **Section 7.3**.

Modified Intent-to-Treat (mITT)

The mITT population includes all randomized subjects who satisfy all major eligibility criteria.

Failure to meet the following eligibility criteria will result in exclusion from the mITT population:

Inclusion:

- A standing radiograph of the knee showing a Kellgren-Lawrence grade of 2 to 4 and an absence of severe osteoarthritis (defined as advanced stage osteoarthritis, including large osteophytes, chronic fractures or bone remodeling, severe deformity or bone attrition, and/or bone-on-bone contact indicative of severe osteoarthritis/ full thickness cartilage loss), as confirmed by the central imaging laboratory.
- A WOMAC LK 3.1 pain subscale total score ≥ 9 and ≤ 19.
- Signed an institutional review board approved informed consent.

Exclusion:

- Presence of symptomatic osteoarthritis in the non-study knee; if unclear then the WOMAC LK 3.1 pain sub-scale for the non-index knee must be ≤ 5.0.
- Diagnosed with rheumatoid arthritis, Reiter's syndrome, psoriatic arthritis, gout, ankylosing spondylitis, or arthritis secondary to other inflammatory diseases; HIV, viral hepatitis; chondrocalcinosis, Paget's disease, or villonodular synovitis.

In addition, subjects may be withdrawn by the investigator prior to receiving an injection for the following reasons:

- If it is not possible to withdraw the required amount of blood for nSTRIDE APS Kit processing (60 milliliters), then the subject will be withdrawn from the study and will not receive the injection. This will ensure that all treated subjects receive the appropriate treatment volume.

- If, on the day of injection, it is determined by the investigator that the subject is no longer an appropriate candidate for injection (e.g., due to development of infection of the index joint) then the subject will be withdrawn from the study and will not receive the injection.

Subjects who are withdrawn from the study prior to receiving an injection will be replaced and will not be included in the mITT population.

This population will be used as a sensitivity analysis of the primary endpoint.

Per Protocol (PP)

The PP population includes all subjects from the mITT population who do not have major protocol deviations that impact patient safety or the scientific validity of the study outcomes, or the potential to do so. Analyses which use the PP population will use all available data on these subjects, with no imputation of missing data. This is the population that will be used in the analysis of secondary and exploratory efficacy endpoints, and as a sensitivity analysis of the primary endpoint.

6.4 Subjects with Missing Data

The primary analysis will be performed on the ITT population. In the primary analysis, missing WOMAC Pain scores for subjects who are withdrawn due to reasons related to the index knee OA ("non-responders") will be considered Missing Not At Random (MNAR), and the subjects' baseline score will be carried forward in the analysis of WOMAC Pain. Scores missing for all other reasons will be considered Missing At Random and will be imputed using a Multiple Imputation method, which will replace each missing value with a set of plausible values that represent the variability around the choice of which value to impute. The monotone[1] regression method will be used [23], and will include variables from **Table 1** (i.e. gender, age, BMI, race, and baseline WOMAC Pain) as well as treatment group. As suggested in Bodner and White et al. [24, 25], the number of imputations will reflect the percentage of incomplete cases. Approximately 15% attrition is estimated in the sample size calculation and thus approximately fifteen imputed datasets will be incorporated. A seed of 20180101 will be used in generating the random numbers.

No imputation of missing values is planned for secondary or exploratory outcome variables.

6.5 Sensitivity Analysis

Scores for subjects who are withdrawn due to reasons related to the index knee OA are considered Missing Not At Random. In all of the ITT analyses

[1] If data are severely non-normal based on a graphical review, then the predictive mean matching method will be used instead of the monotone regression method.

described below, these subjects will be considered “non-responders”, and the subjects’ baseline score will be used in the analysis of WOMAC Pain.

The primary study analysis assumes that the all other missing data is Missing At Random (MAR).

In order to investigate the assumption that the missing data (excluding the non-responders who are withdrawn for reasons related to the index knee OA) is MAR, sensitivity analysis of the primary study endpoint will be performed under assumption of other missing data mechanisms, as follows:

1. A complete case analysis on the ITT population (Missing Completely at Random (MCAR))
2. A complete case analysis on the Per Protocol population (Missing Completely at Random (MCAR))
3. A tipping point analysis on the ITT population, performed under the assumption that the data is not MAR or MCAR by searching for a tipping point that reverses the study conclusion. This analysis will examine the possibility that the distribution of missing responses will have a different expected value than that of the corresponding distribution of the observed responses. Thus, the analysis will generate multiple imputed data sets with a specified sequence of shift parameters that adjust the imputed values for observations in the treatment group (Missing Not At Random (MNAR)).

In the event the conclusion from one or more of the sensitivity analyses disagrees with the conclusion from the primary analysis, the sources of differences between them will be investigated and subjected to explicit discussion and interpretation.

Prior results indicate that the efficacy of APS and the saline control differs over time, but this difference is small until some point after the 6 month post-injection time point. Therefore, although carrying data forward from an earlier time is commonly applied as a sensitivity analysis, in this case it is likely to produce misleading results if applied to primary endpoints and so is not planned.

7. Data Collection and Reporting

7.1 Data Collection Time Points

Data will be collected at the screening visit, the injection visit, and at 1, 3, 6 and 12 months following the injection. At 12 months, after the 12 month evaluation is complete, subjects will be un-blinded. At this time the subject may opt to get an injection of APS (regardless of original treatment assignment). If a second injection occurs, then a follow-up visit will be scheduled, and data will be captured approximately 1 month after the second

injection. A table showing the data collected at each time interval is presented in the study protocol.

7.2 Principal Data Collected

The principal data collected will be the WOMAC LK 3.1 Index. The primary endpoint is the WOMAC Pain score. WOMAC Function will be used as a secondary endpoint, and WOMAC Stiffness and Total WOMAC will be used as exploratory endpoints.

In addition, VAS Pain and the EQ-5D questionnaires will be administered.

A pre-procedure X-ray and MRI will be performed and then repeated at 12 months post-procedure.

Blood, APS and synovial fluid (where aspiration is productive) will be sent to a central laboratory. Blood and APS specimens will be characterized with respect to cell counts, cytokine quantification and growth factor composition as appropriate for the specimen type.

Acetaminophen and restricted medication use, as well as adverse events and protocol deviations, will be recorded over the course of the study.

7.3 Data Reporting

Descriptive statistics will be provided for each demographic and outcome variable collected in the study, in tabular form by treatment group. For interval level variables, at a minimum, this will include mean, maximum, minimum and standard deviation. Categorical data will be displayed including counts and frequency percent by treatment group. Boxplots including standard error bars may also be presented for continuous endpoints.

8. Summary of Data Analyses

Table 10 summarizes the planned analyses that are laid out in this report.

Table 10. Planned Analyses

Table Number	Table Title
Table 1	Randomization Verification
Table 3a	Planned Primary Efficacy Test
Table 3b	Planned Secondary Efficacy Tests
Table 3c	Planned Exploratory Efficacy Tests: Questionnaire Data
Table 4a	Planned Statistical Tests for X-ray Results

Table 4b	Planned Statistical Tests for MRI Results
Table 4c	Planned Statistics for Imaging/ Questionnaire Results
Table 5a	Planned Statistical Analyses for APS / Blood Results (All-inclusive and Anti-inflammatory)
Table 5b	Planned Statistical Analyses for APS / Blood Results (WBCs and Growth Factors)
Table 5c	Planned Statistical Analyses Comparing APS / Blood Results
Table 6	Subgroup Analysis
Table 7	Analysis of Data after 2nd Injection

9. References

1. Brzusek, D. and D. Petron, *Treating knee osteoarthritis with intra-articular hyaluronans.* Curr Med Res Opin, 2008. **24**(12): p. 3307-3322.
2. Goldring, S.R. and M.B. Goldring, *The role of cytokines in cartilage matrix degeneration in osteoarthritis.* Clin Orthop Relat Res, 2004(427 Suppl): p. S27-S36.
3. Goldring, M.B., *Anticytokine therapy for osteoarthritis.* Expert Opin. Biol Ther, 2001. **1**(5): p. 817-829.
4. Goldring, M.B., et al. *Defining the roles of inflammatory and anabolic cytokines in cartilage metabolism*. in *Annals of the Rheumatic Diseases*. 2008.
5. Goldring, M.B., *Update on the biology of the chondrocyte and new approaches to treating cartilage diseases.* Best Pract. Res Clin Rheumatol, 2006. **20**(5): p. 1003-1025.
6. O'Shaughnessey, K., et al., *Autologous protein solution prepared from the blood of osteoarthritic patients contains an enhanced profile of anti-inflammatory cytokines and anabolic growth factors.* J Orthop Res, 2014. **32**(10): p. 1349-1355.
7. Dinarello, C.A., *The role of the interleukin-1-receptor antagonist in blocking inflammation mediated by interleukin-1.* New England Journal of Medicine, 2000. **343**(10): p. 732-734.
8. Van Zee, K.J., et al., *Tumor necrosis factor soluble receptors circulate during experimental and clinical inflammation and can protect against excessive tumor necrosis factor alpha in vitro and in vivo.* Proc Natl. Acad Sci U. S. A, 1992. **89**(11): p. 4845-4849.
9. O'Shaughnessey, K., A. Panitch, and J.E. Woodell-May, *Blood-derived anti-inflammatory protein solution blocks the effect of IL-1b on human macrophages in vitro.* Inflamm. Res, 2011. **60**(10): p. 929-936.
10. Matuska, A., et al., *Autologous solution protects bovine cartilage explants from IL-1a and TNFa induced cartilage degradation.* Orthopaedic Research Society, 2013. **31**(12): p. 1929-1935.
11. Dmitrienko, A., et al., *Analysis of clinical trials using SAS: A practical guide*. The Power To Know. 2005, Cary, NC 27513: SAS Institute.
12. Maurer, W., *Multiple comparisons in drug clinical trials and preclinical assays: a-priori ordered hypothesis.* Biomed. Chem.-Pharm. Ind., 1995. **6**: p. 3-18.
13. Bannuru, e.a., *Did the American Academy of Orthopaedic Surgeons Miss the Mark?* Arthroscopy: The Journal of Arthroscopic and Related Surgery, 2014: p. pp 86-89.
14. *Methods to Identify What is Important to Patients & Select, Develop or Modify Fit-for-Purpose Clinical Outcomes Assessments*. 2018.
15. Bellamy, *Towards a definition of "difference" in osteoarthritis.* J Rheumatol, 2001: p. 8:427-430.

16. Guyatt, G.H., et al. *Methods to explain the clinical significance of health status measures*. in *Mayo Clinic Proceedings*. 2002. Elsevier.
17. Jevsevar, D.S., *Treatment of osteoarthritis of the knee: evidence-based guideline.* JAAOS-Journal of the American Academy of Orthopaedic Surgeons, 2013. **21**(9): p. 571-576.
18. Pham, T., et al., *OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited.* Osteoarthritis Cartilage, 2004. **12**(5): p. 389-399.
19. Jang, S.J., J.D. Kim, and S.S. Cha, *Platelet-rich plasma (PRP) injections as an effective treatment for early osteoarthritis.* European Journal of Orthopaedic Surgery & Traumatology, 2012.
20. Cerza, F., et al., *Comparison Between Hyaluronic Acid and Platelet-Rich Plasma, Intra-articular Infiltration in the Treatment of Gonarthrosis.* American Journal of Sports Medicine, 2012.
21. Bolognese, J., T. Schnitzer, and E. Ehrich, *Response relationship of VAS and Likert scales in osteoarthritis efficacy measurement.* Osteoarthritis and cartilage, 2003. **11**(7): p. 499-507.
22. CDRH, *Guidance for Industry and FDA Staff - Clinical Data Presentations for Orthopedic Device Applications*, F.a.D.A. U.S. Department of Health and Human Services, Center for Devices and Radiological Health, Orthopedic Devices Branch, Division of General, Restorative, and Neurological Devices, Office of Device Evalution, Editor. 2004.
23. Rubin, D.B., *Multiple imputation for nonresponse in surveys*. Vol. 81. 2004: John Wiley & Sons.
24. Bodner, T.E., *What improves with increased missing data imputations?* Structural Equation Modeling: A Multidisciplinary Journal, 2008. **15**(4): p. 651-675.
25. White, I.R., P. Royston, and A.M. Wood, *Multiple imputation using chained equations: issues and guidance for practice.* Statistics in medicine, 2011. **30**(4): p. 377-399.

Statistical Analysis Plan Revision History		
SAP Version	**Description of Change**	**Sections**
V 1.0	Original SAP submitted to FDA (Dated 2016-06-28)	
V 2.0*	Clarified how Type I error is handled in the primary analysis by setting up a pre-specified, fixed sequential order of testing.	Study Objectives (Section 1) Primary Outcomes (Section 4.1)
	Clarified how Type I error is handled in the secondary analysis by setting up a pre-specified, fixed sequential order of testing.	Study Objectives (Section 1) Secondary Outcomes (Section 4.2)
	Added repeat of primary efficacy tests in Kellgren-Lawrence Grade II Subgroup to planned secondary endpoints	Secondary Outcomes (Section 4.2)
	Changed some endpoints previously specified as “Secondary” to “Exploratory” endpoints.	Study Objectives (Section 1) Exploratory Outcomes (Sections 4.3-4.5)
	Moved description of imaging outcomes from the “Study Objectives” section to separate sections (Exploratory Outcomes (Imaging Outcomes)).	Study Objectives (Section 1) Exploratory Outcomes (Imaging Outcomes) (Section 4.4)
	Moved description of laboratory outcomes from the “Study Objectives” section to separate section (Exploratory Outcomes (Laboratory Outcomes)).	Study Objectives (Section 1) Exploratory Outcomes (Laboratory Outcomes) (Section 4.5)
	Moved description of analysis populations from the “Study Design” section to “Analysis Populations Assignment” section.	Study Design (Section 1.2) Analysis Populations Assignment (Section 6.3)
	Created a separate section for “Analysis of Baseline Data”. Clarified testing to confirm equivalence of treatment groups at baseline; removed MANOVA and WBC Count variable to simplify the testing.	Analysis of Baseline Data (Section 2)

	Created a separate section for “Subject Disposition”.	Subject Disposition (Section 3)
	Added a table to clarify how visit windows would be calculated using “days since injection” (Table 2, Visit Windows”).	Subject Disposition (Section 3)
	Separated “Outcomes” section into “Effectiveness Outcomes” and “Safety Outcomes” sections.	Effectiveness Outcomes (Section 4) Safety Outcomes (Section 5)
	Changed the method of evaluating normality of continuous effectiveness outcomes; they will be evaluated graphically instead of using the Shapiro-Wilk test.	Effectiveness Outcomes (Section 4)
	Described and justified fixed-sequential testing to be used for primary analysis; added additional reference.	Primary Outcomes (Section 4.1)
	Clarified conclusions to be made if superiority is found with regard to (1) WOMAC Pain subscale only, or (2) both WOMAC Pain and Function subscales.	Primary Outcomes (Section 4.1)
	Added “Order of Testing” to the table describing the Primary Efficacy tests and expected outcomes.	Primary Outcomes (Section 4.1)
	Described and justified fixed-sequential testing to be used for secondary analyses.	Secondary Outcomes (Section 4.2)
	Justified the chosen order of testing of secondary endpoints.	Secondary Outcomes (Section 4.2)
	Added “Order of Testing” to the table describing the Secondary Efficacy tests and expected outcomes.	Secondary Outcomes (Section 4.2)
	Changed all variables previously designated as “secondary” to “exploratory”, with exception of the change in VAS Pain and OMERACT-OARSI responder outcomes.	Secondary Outcomes (Section 4.2) Exploratory Outcomes (Questionnaire) (Section 4.3)
	Separated descriptions of Exploratory Analyses into sections for Questionnaire Outcomes (Section 4.3), Imaging Outcomes (Section 4.4), and Laboratory Outcomes (Section 4.5).	Exploratory Outcomes (Questionnaire) (Section 4.3) Exploratory Outcomes (Imaging) (Section 4.4) Exploratory Outcomes (Laboratory)

		(Section 4.5)
	Clarified that Exploratory Outcomes would be tested at alpha = 0.05 with no adjustment for multiple comparisons.	Exploratory Outcomes (Questionnaire) (Section 4.3) Exploratory Outcomes (Imaging) (Section 4.4) Exploratory Outcomes (Laboratory) (Section 4.5)
	In the Planned Exploratory Outcomes (Table 3c), separated EQ-5D change score from EQ-5D dimensions, and added appropriate statistical tests.	Exploratory Outcomes (Questionnaire) (Section 4.3)
	Added detail about the Imaging outcomes and associated statistical tests to be used to analyze them. Updated and added some statistical testing to align with planned imaging data collection and derivation activities, as defined by the Image Review Charter for this study.	Exploratory Outcomes (Imaging) (Section 4.4)
	Added detail about the Laboratory outcomes, planned testing, and associated statistical tests to be used. Updated some statistical testing to be used to make more appropriate for the associated endpoint.	Exploratory Outcomes (Laboratory) (Section 4.5)
	Changed method of evaluating normality of continuous laboratory outcomes; they will be evaluated graphically instead of using the Shapiro-Wilk test. Updated some statistical testing to be used to make more appropriate for the associated endpoint.	Exploratory Outcomes (Laboratory) (Section 4.5)
	Added sentence to describe analysis of use of Rescue Medication.	Safety Outcomes (Section 5)
	Sample Size details were updated to: (1) Ensure that study is powered adequately for both WOMAC Pain and Function subscales, (2) Adjust for attrition of only 15% instead of 20% as before.	Sample Size (Section 6.1)
	Added a section describing the different study populations for analysis.	Analysis Populations Assignment (Section 6.3)

	Clarified sensitivity analyses descriptions, to specify that (1) sensitivity analyses will be performed regardless of the results of the baseline comparisons, and (2) sensitivity analyses will be conducted on the primary endpoints only.	Subjects with Missing Data (Section 6.4)
	Made minor clarifications to specify the principal data collected.	Principal Data Collected (Section 7.1)
	Made minor clarifications to specify plans for data reporting.	Data Reporting (Section 7.2)
	Made minor editorial clarification changes throughout.	All Sections
V 3.0	Updated description of secondary analyses to match currently planned analyses and ranked order of analyses as defined in Section 4.2	Introduction (Section 1)
	Updated planned sample size per Section 6.1	Introduction (Section 1.2)
	Clarified purpose of baseline comparisons.	Analysis of Baseline Data (Section 2)
	Added optional visits	Subject Disposition (Section 3)
	Specified summaries of concomitant medications and comorbidities as part of baseline data analysis.	Analysis of Baseline Data (Section 2)
	Added figure 1 to clarify primary analyses	Primary Outcomes (Section 4.1)
	Specified type of fixed-sequence procedure to be performed on primary outcomes	Primary Outcomes (Section 4.1)
	Clarified that any analyses conducted after a sequential null hypothesis failed to be rejected would be conducted as exploratory only.	Secondary Outcomes (Section 4.2)
	Moved VAS Pain to last secondary analysis, since it is highly correlated with WOMAC pain, which is included in the primary analysis.	Secondary Outcomes (Section 4.2)
	Specified OMERACT-OARSI criteria	Secondary Outcomes (Section 4.2)
	Added WOMAC Function as a dependent variable in analysis of clinical and MRI outcomes	MRI Outcomes (Section 4.4.2)
	In the Planned Exploratory Outcomes (Table 3c), added Percent change analyses and within-group analysis over time.	Exploratory Outcomes (Questionnaire) (Section 4.3)
	Added assessment of key APS output characterization criteria and plan for listing any samples not meeting the criteria to review the clinical data and evaluate any relationship between the characterization criteria and clinical outcomes	Exploratory Outcomes (Laboratory Tests of Blood and APS) (Section 4.5)
	Specified subgroup analyses to be conducted per recommendations in “Evaluation and Reporting of Age-, Race-, and Ethnicity-Specific Data in Medical Device Clinical Studies.”	Subgroup Analyses (Section 4.6)
	Added Section 4.7 to describe analysis of data post-2nd Injection	Analysis of Data after 2nd Injection

		(Section 4.7)
	Added reference to protocol section for definitions of AE, SAE, ADE, and SADE. Modified time-course table so that AE windows are specified without specifying a specific format for the shell table.	Adverse Events (Section 5.1)
	Specified analysis measures to account for use of Rescue Medication, per FDA Statistical Considerations from IDE approval letter.	Rescue Medication (Section 5.2)
	Specified planned imputation of primary outcomes for subjects who are withdrawn due to reasons related to the index knee OA will be considered "non-responders",	Rescue Medication (Section 5.2)
	Added Kaplan Meier analysis for time to "non-response" (withdrawal due to reasons related to the index knee)	Rescue Medication (Section 5.2)
	Adjusted sample size to provide a minimum of 90% power	Sample Size (Section 6.1)
	Added reference for attrition estimate used in the sample size calculation.	Sample Size (Section 6.1)
	Removed description of blinding, since detail is already included in blinding plan.	Treatment Assignments (Section 6.2)
	Changed population to be used for primary analysis from PP to mITT, per FDA Statistical Considerations from IDE approval letter.	Analysis Populations Assignment (Section 6.3)
	Stated that list of subjects with major protocol deviations that impact patient safety or the scientific validity of the study comparisons, or have the potential to do so, will be performed by a blinded review of the data.	Analysis Populations Assignment (Section 6.3)
	Defined ITT, mITT and PP populations, including a predefined list of "major" eligibility violations and criteria for study withdrawal prior to injection.	Analysis Populations Assignment (Section 6.3)
	Defined mechanism for accounting for missing data in the primary analysis.	Subjects with Missing Data (Section 6.4)
	Defining sensitivity analyses, as well as what will take place if the outcome of primary and sensitivity analyses disagree.	Sensitivity Analysis (Section 6.5)
V 3.1	Redefined primary endpoint to include WOMAC Pain only, with WOMAC Function defined as a secondary endpoint. Redefined the indication for surgery to include "Pain associated with OA".	Study Objectives (Section 1), Effectiveness Outcomes (Section 4), Data Collection and Reporting (Section 7)
	Updated the variable Ethnicity to Race, to reflect the CRF and other references to this variable in the SAP.	Analysis of Baseline Data (Section 2)
	Removed Baseline (Screening) WOMAC Function from the analyses designed to evaluate baseline characteristics across treatment groups.	Analysis of Baseline Data (Section 2)
	Modified Planned Primary Efficacy Test table (Table 3a) to remove Order of Testing column, since only one test is being performed for the new primary endpoint.	Primary Outcomes (Section 4.1)
	Removed flowchart Indications for Use (Figure 1) and text reference to flowchart, since the new primary endpoint only allows for one indications for use option.	Primary Outcomes (Section 4.1)

	Renamed Figure 2 to Figure 1, subsequent to deletion of Figure 1 in Section 4.1.	Secondary Outcomes (Section 4.2)
	Indicated that the sample size was powered for both the updated primary endpoint (WOMAC Pain) and the updated first secondary endpoint (WOMAC Function).	Study Populations (Section 6)
	Designated the ITT population as the population to be used in the primary study analysis.	Study Populations (Section 6)
	Changed Sponsor representative from Krista Toler to Ann Blanton.	Signature Page
V 4.0	Redefined primary endpoint to include mean improvement in WOMAC Pain and MCII Responders for Function as sequential primary endpoints, with the indication to include “treatment of knee osteoarthritis and associated symptoms or knee pain associated with osteoarthritis”. Redefined secondary endpoint to include mean improvement in WOMAC Function.	Study Objectives (Section 1)
	Redefined indication to “treatment of osteoarthritis and associated symptoms or knee pain associated with osteoarthritis” Addition of WOMAC Function to Randomization Verification.	Study Design (Section 1.2)
	Add WOMAC Function to the list of baseline items to be evaluated	Analysis of Baseline Data (Section 2)
	Specify the type of effectiveness outcomes (i.e. continuous)	Effectiveness Outcomes (Section 4)
	Redefined primary endpoint to include mean improvement in WOMAC Pain and MCII Responders for Function as sequential primary endpoints, with the indication for surgery to include “treatment of knee osteoarthritis and associated symptoms” if both endpoints are met, and “knee pain associated with OA” if only pain endpoint is met. Define the MCII responder criteria and primary hypothesis and associated test. Add justification for fixed sequential testing. Planned primary efficacy test for WOMAC Function MCII added to Table 3a. Updated Indication for Use in Figure 1 Flow Chart.	Primary Outcome (Section 4.1)
	Add the assessment of MCII Responder for Function over time to the list of exploratory analyses	Exploratory Outcomes Questionnaire (Section 4.3)
	Add the assessment of relationship between Imaging outcomes and MCII Responder for Function	Exploratory Outcomes Imaging (Sections 4.4.1 and 4.4.2)
	Add the assessment of relationship between laboratory measurements and MCII Responder for Function	Exploratory Outcomes Laboratory Tests of Blood and APS

		(Section 4.5)
	Add analysis of subgroups for MCII Responder for Function	Subgroup Analyses (Section 4.6)
	Updates to Sample Size description to show that the number of enrolled subjects gives sufficient power for updated primary hypothesis	Sample Size (Section 6.1)
	Add missing data consideration for WOMAC Function along with Pain	Subjects with Missing Data (Section 6.4)
	Add WOMAC Function into the Sensitivity analysis considerations	Sensitivity Analysis (Section 6.5)
V 5.0	Redefined primary endpoint to include WOMAC Pain only, with WOMAC Function MCID responder rate defined as a first secondary endpoint. Added descriptive measures to ensure that WOMAC Function increases or remains neutral. Redefined proposed indication as “knee pain associated with osteoarthritis”.	Study Objectives (Section 1), Effectiveness Outcomes (Section 4), Study Populations (Section 6)
	Added rationale regarding the MCID Function criteria	Secondary Outcomes (Section 4.2)
	Added criteria to the OMERACT-OARSI Responder Criteria at the 12 month visit; i.e. no rescue medication between 6-12 months.	Secondary Outcomes (Section 4.2)
	Moved Repeated Measures analysis of WOMAC Pain to the secondary analyses, in accordance with the FDA Analgesics Indications guidance document (https://www.fda.gov/downloads/drugs/guidancecompliancereg ulatoryinformation/guidances/ucm384691.pdf).	Secondary Outcomes (Section 4.2)
	Added 2 additional secondary analyses regarding pain medication usage, in accordance with the FDA Analgesics Indications guidance document (https://www.fda.gov/downloads/drugs/guidancecompliancereg ulatoryinformation/guidances/ucm384691.pdf).	Secondary Outcomes (Section 4.2)
	Added analysis of impact of pain medication usage on improvement in WOMAC Pain score, in accordance with the FDA Analgesics Indications guidance document (https://www.fda.gov/downloads/drugs/guidancecompliancereg ulatoryinformation/guidances/ucm384691.pdf).	Exploratory Outcomes (Section 4.2.1)
	Added information to show that study is sufficiently powered to test the primary analysis and first two secondary analyses.	Sample Size (Section 6.1)
	Added “boxplots with standard error bars” as a descriptive statistic for continuous variables.	Data Reporting (Section 7.3)
	Added Section 8 to summarize all statistical analyses in the document.	Summary of Data Analyses (Section 8)
V 6.0	Added secondary endpoint of MCID Pain responders	Secondary Outcomes (Section 4.2)
	Added updated rationale regarding the MCID Function criteria per FDA comments 30MAY2019	Secondary Outcomes (Section 4.2)
	Added rationale regarding the MCID Pain criteria	Secondary Outcomes (Section 4.2)
	Updated the MCID value and hypothesis test statement for WOMAC Function; added MCID Value and hypothesis test statement for WOMAC Pain	Secondary Outcomes (Section 4.2)

	Added separate analyses for the effects of restricted and rescue medications on the primary outcome, separately for OA and any reasons.	Secondary Outcomes (Section 4.2)
	Added additional exploratory endpoint of MCID Pain responders at 1, 3, and 6 months	Exploratory Outcomes (Section 4.3)
	Added separate analyses for the effects of restricted and rescue medications on the improvement in WOMAC Function, separately for OA and any reasons.	Exploratory Outcomes (Section 4.3)
	Added footnote to clarify the OMERACT-OARSI criteria being used	Exploratory Outcomes (Section 4.3)
	Added a subgroup analysis for the MCID Pain responder endpoint	Subgroup Analyses (Section 4.6)
	Specified baseline covariates for the MI model; increased the number of imputations per Bodner and White, et al.	Subjects with Missing Data (Section 6.4)
V 7.0	Specified an ANCOVA model to analyze trends in Δ WOMAC Pain for treatment / Crossover groups.	Exploratory Outcomes (Section 4.3)
	Added an additional analysis of the change in K-L grade by baseline K-L grade	Exploratory Outcomes (Section 4.4)
	Added additional subgroups for analysis; specified that LS Means would be examined for each subgroup.	Subgroup Analysis (Section 4.6)

*V 2.0 and V 6.0 were signed but not implemented

Signature Page

Statistical Analysis Plan
For APSS-44-00

Protocol Title:

A Multicenter, Double-Blind, Randomized, Saline-Controlled Study of a Single, Intra-Articular Injection of Autologous Protein Solution in Patients with Knee Osteoarthritis

Ann Blanton (Nov 7, 2019)

Ann Blanton Date
Clinical Project Lead, Biologics (Study Manager)
Zimmer Biomet, Inc.
Sponsor Representative

Katie Miller (Nov 7, 2019)

Katie Miller Date
Senior Biostatistician
M Squared Associates, Inc.
Study Statistician