Statistical Analysis Plan (SAP)

A Pivotal Study Comparing Two Injections of MONOVISC to Two Injections of Saline in Patients with Osteoarthritis of the Hip

SHORT TITLE: The MONOVISC Hip OA IDE Study
Protocol Number: 15-MVH-01
Protocol Version: 3 (Amendment 2)
Protocol Revision Date: 01-MAR-2016

SAP Version: 1.0
SAP Version Date: 16-JUN-2016

A Pivotal Study Comparing Two Injections of MONOVISC to Two Injections of Saline in Patients with Osteoarthritis of the Hip

Protocol Version: 3

The following individuals have reviewed this version of the Statistical Analysis Plan and are in agreement with the content:

[redacted]

Revision History

Revision Number	Revision Date (DDMMMYYYY)	Reasons for Revision
1.0	**16JUN2016**	**Initial version**

Table of Contents

1 List of Abbreviations

AE	Adverse Event
AT	As Treated
BSDM	Biostatistics and Data Management
CSP	Clinical Study Protocol
IA	Interim Analysis
IDE	Investigational Device Exemption
ITT	Intent-to-Treat
H_A	Alternative Hypothesis
H_O	Null Hypothesis
N/A	Not Applicable
NRS	Numerical Rating Scale
MITT	Modified Intent to Treat
OA	Osteoarthritis
PGA	Patient Global Assessment
PP	Per-Protocol
SAP	Statistical Analysis Plan
WOMAC	Western Ontario and McMaster Osteoarthritis Index

2 Modifications in Analysis from the Clinical Study Protocol

The basis of this SAP is the Clinical Study Protocol (CSP). For reasons of clarity certain redundant text has been omitted and changes have been made to address any typographical errors and improve consistency of terminology. In addition, substantive changes are presented in Table 1, below.

Table 1 Summary of Substantive Changes

Content	Modification	Justification
Block Randomization, stratified by study site	One Master Randomization List was create with which to sequentially label patient kits as MONOVISC or saline.	Stratification of randomization by study site was not feasible because it would have required a significant quantity of additional product, storage of randomized product by site, and added complexity in procedures with the supply vendor.
Mean Change from Baseline through 26 Weeks in WOMAC A1	Added endpoint and method of analysis	In addition to assessing the primary endpoint, WOMAC A1 walking pain

Content	Modification	Justification
		score change from baseline at 26 weeks, interest lies in assessing the difference in the mean change from baseline in WOMAC A1 over the entire study period.
Interim Analysis	An unblinded analysis team (separate and firewalled from the study team) will be responsible for the conduct of the Interm Analysis rather than a member of the Strategic Medical Affairs department	As the analysis will be used to make a go/no go decision on study conduct, a rigorous approach to the analysis is warranted.

3 The Study Objectives

The primary objective of this study is determine whether two intra-articular injections of MONOVISC, separated by 1 month, are superior to two intra-articular injections of physiologic saline, separated by 1 month, in relieving hip osteoarthritis pain, as determined by reduction in walking pain change from baseline.

We hypothesize that two (2) monthly injections of MONOVISC will be an effective treatment regimen for hip OA patients.

4 Study Design

This study is a prospective, multi-center, double-blinded, randomized, controlled, superiority study comparing intra-articular injections of MONOVISC High Molecular Weight Hyaluronan with intra-articular injections of physiologic saline. The study will take place in the United States, with up to 25 investigational sites. Subjects will be randomized through a 2:1 schema to receive either two (2) injections of 4 ml of MONOVISC, separated by 1 month (active treatment), or two (2) monthly injections of 4 ml of saline, separated by 1 month (control treatment).

Subjects will return for follow-up visits at 2, 4, 8, 16, and 26 weeks after the first injection.

The primary endpoint will be pain during walking (WOMAC A1) change from baseline at the 26 week time point.

An interim analysis (IA) of the primary endpoint will be conducted after 74 MONOVISC and 37 saline subjects have received both injections of MONOVISC or saline and completed their per protocol 26 week follow up visit. A stopping rule for futility will be implemented, such that the study will be terminated if the improvement of MONOVISC over saline (change from baseline) is less than 0.4 on a 10 point Numerical Rating Scale (NRS) in both the Per Protocol Analysis Set for the 26 week interval and the As Treated Analysis Set for the 26 Week interval. If the improvement of MONOVISC over saline (change from baseline) is more than 0.4 NRS in either of these analysis sets, the study will continue to completion. Details of planned analyses are provided in Section 14.

5 Treatment Assignment

The treatment groups in this study are:

MONOVISC: Two intra-articular injections of MONOVISC High Molecular Weight Hyaluronan, separated by 1 month (active treatment).

Saline: Two injections of 4 ml saline, separated by 1 month (control treatment).

5.1 Randomization

Subjects will be randomized in a 2:1 ratio to receive either treatment with MONOVISC or treatment with Saline. Block randomization (random blocks of size 3 or 6) will be used to create a Master Randomization List and patient kits of MONOVISC and saline will be sequentially numbered according to this list. With this randomization scheme, each study subject will be associated with a unique kit number, and the order of randomized subjects at each investigational site will be distinct. Prior to initiation of the study, investigational sites will be provided with an initial, sequentially numbered quantity of kits to be used, and as these kits are utilized, additional quantities of sequentially numbered kits can be ordered as needed through an electronic re-order system. The kit will not contain any labeling to indicate whether the patient is receiving MONOVISC or Saline. After the first injection, the remaining syringe will be returned to the storage site for the second injection.

5.2 Blinding

All clinical research and biostatistics personnel at DePuy Synthes Mitek Sports Medicine will remain blinded to treatment assignment until the end of the study with 2 exceptions:

- an employee within DePuy Synthes Mitek Sports Medicine Strategic Medical Affairs will be designated to maintain the Master Randomization, to facilitate any medical safety reviews, if necessary, and to facilitate the planned interim futility analysis.
    - Rodrigo Diaz, Franchise Medical Director, DePuy Synthes is currently fulfilling this role. If a change is required, a replacement for this responsibility will be assigned and documented in a note to file.
- A team of qualified employees from within DePuy Synthes Mitek Sports Medicine Biostatistics and Data Management (BSDM) department who will be designated to conduct the Interim Analysis. Members of this team will not be otherwise associated with the conduct of the study.

6 Levels of Significance

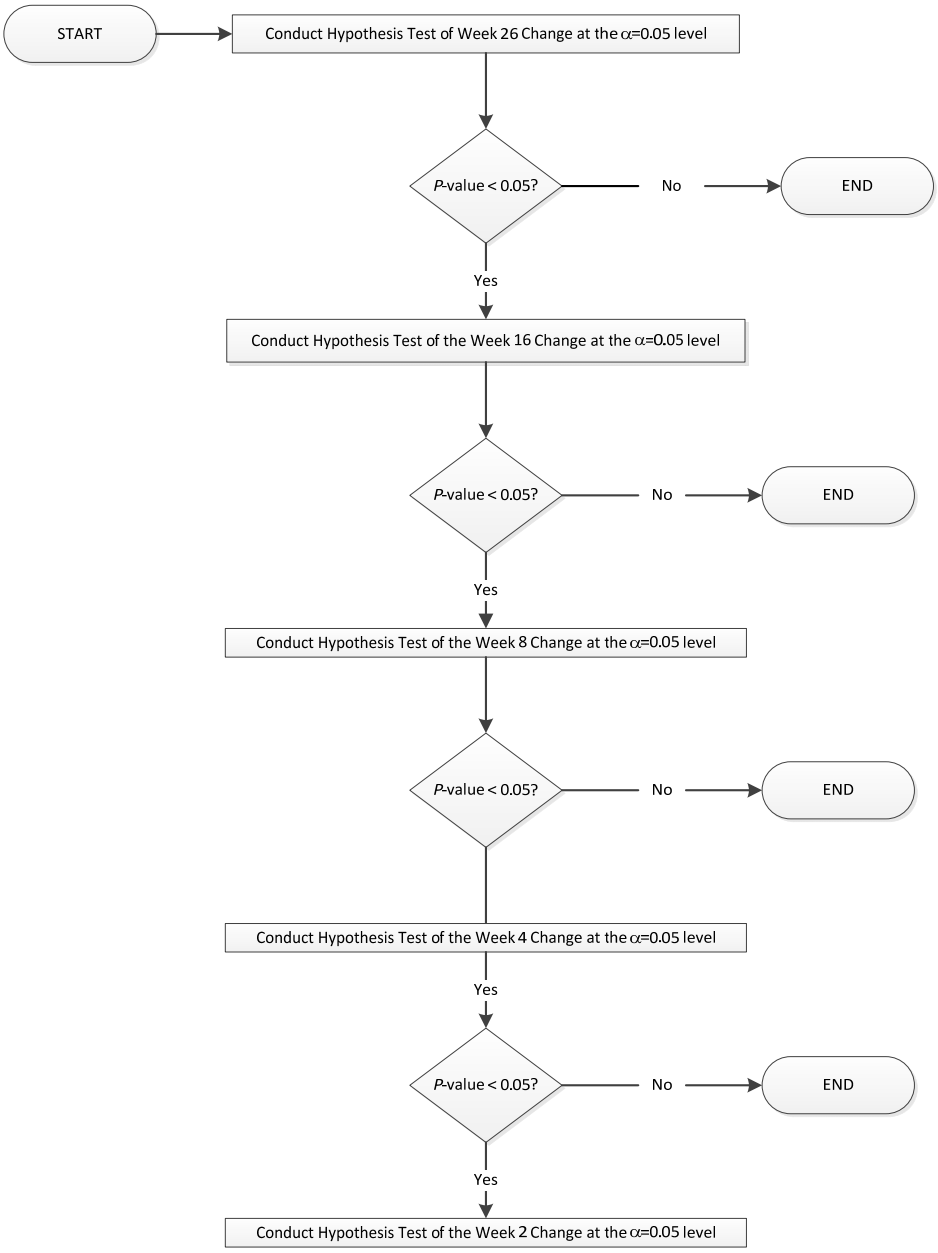
In general, confidence intervals will be 2-sided, 95% confidence intervals, and p-values will also be 2-sided; p-values below 0.05 will be deemed to be statistically significant.

6.1 Control of Type 1 Error

Familiy-wise control of Type I Error across all primary and secondary endpoints will be accomplished using the following gatekeeping strategy which is depicted in Figure 1.

If the study is successful and the primary endpoint analysis demonstrates a statistically significant difference (favoring MONOVISC vs. Saline) at 26 Weeks, then we will test the WOMAC A1 change from baseline at 16 Weeks for a significant difference (favoring MONOVISC vs. Saline). If this is successful, we will test the difference at 8 Weeks; if this is successful, we will test the difference at 4 Weeks; and if this is successful, we will test the difference at 2 Weeks. At any point, if one of these comparisons fails to be statistically significant, then we will cease further testing in the specified order. Each of these tests will be conducted utilizing the primary endpoint analysis on the Modified Intent to Treat Analysis Set for Longitudinal Analysis. We will test each of the secondary endpoints with a 2-sided p-value against threshold of 0.05 (a 2-sided 5% alpha) with no further adjustment for multiplicity. Further detail of the hypotheses tested is found in Section 10.

Figure 1 Primary and Secondary Gatekeeping Procedural Flow



Note: The hypothesis tests indicated above are to be performed on the endpoints which we have specified as primary and secondary, WOMAC A1 change from baseline at 26 (primary) weeks and also at 2, 4, 8, and 16 weeks.

7 Interval Windows

The study windows are presented in the Time and Events Schedule (CSP Section 1.1).

For analysis purposes, the baseline/screening evaluation is to be done within 30 days prior to the first injection; the first injection is considered to be day 0. The 2, 4, 8, 16 and 26 Week visits will occur in the time intervals (days) indicated in Table 2, below.

Table 2 Study Interval Windows

Analysis Visit Label	Nominal Study Visit	Study Interval
Baseline	Baseline/Screening	-30 to 0
2 Week	Day 14 $\pm$ 7 days	11 to 17
4 Week	Day 28 $\pm$ 7 days	21 to 35
8 Week	Day 60 $\pm$ 14 days	46 to 74
16 Week	Day 120 $\pm$ 14 days	105 to 134
26 Week	Day 180 $\pm$ 14 days	159 to 201

Note:
Day 0 is defined as the Study Day of first injection.
Baseline is defined as the last non-missing assessment prior to first injection.
Baseline must occur no more than 30 days after Screening.

8 Handling of Missing Data

The primary efficacy endpoint analysis will be conducted on the MITT Analysis Set, which consists of all primary endpoint data on all As Treated (AT) subjects (regardless of whether or not there were protocol deviations associated with the data) for the 2, 4, 8, 16 and 26 week visits. In this MITT Analysis Set, missing primary endpoint data for the 26 Week visit will be imputed with multiple imputation methodology using SAS PROC MI with the regression method and a monotone missing data pattern.

No imputation of missing data will be performed for visits other than the 26 Week visit.

9 Primary and Secondary Endpoints

The primary and secondary endpoints are derived from the Western Ontario and McMaster Universities Arthritis Index (WOMAC), which is widely used in the evaluation of Hip and Knee Osteoarthritis. The WOMAC is available in over 65 languages and has been linguistically validated. It is a self-administered questionnaire consisting of 24 items divided into 3 subscales, A, B, and C: A) Pain (5 items), B) Stiffness (2 items), and C) Physical Function (17 items). The overall WOMAC score is determined by summing the scores across the 3 subscales and the score ranges include 0–24 (derived from the NRS scale). Higher scores on

the WOMAC indicate worse pain, stiffness, and functional limitations. Specific scoring details will be provided in the table shells document referenced in Section 15.

The WOMAC A is the pain subscale of the WOMAC questionnaire. Again, response options for individual questions are 0 through 10, where 0 indicates no pain and 10 indicates worst possible pain. The WOMAC A pain subscale scores ranges are 0-50 NRS.

WOMAC A1 is the first question of the WOMAC A and specifically addresses pain during walking. Response options are 0 through 10 NRS, where 0 indicates no pain and 10 indicates worst possible pain.

9.1 Primary Endpoint

The primary efficacy endpoint in this study is the WOMAC A1 walking pain score change from baseline at 26 weeks.

9.2 Secondary Endpoints

The secondary endpoints in this study are also based on the WOMAC A1 walking pain score change from baseline; four secondary endpoints are defined, one for each of the remaining post-baseline time points: at the 2, 4, 8, and 16 Week time points. Analysis will be conducted using the MITT Longitudinal Analysis Set and will include data from weeks 2 through 26 weeks. However, no imputation will be done for any time point other than the 26 Week.

9.3 Tertiary Endpoints

The following are considered tertiary endpoints, and will be summarized in tables of summary statistics:

Walking Pain:

WOMAC A1 score mean change over 26 Weeks.WOMAC A1 score at the 2, 4, 8, 16 and 26 Week time points.

WOMAC A1 score change from baseline area under the curve (AUC).

Pain:

WOMAC A score at the 2, 4, 8, 16 and 26 Week time points.

WOMAC A score change (improvement) from baseline at the 2, 4, 8, 16 and 26 Week time points.

Pain, Stiffness and Physical Functioning:

Total WOMAC score at the 2, 4, 8, 16 and 26 Week time points.

Total WOMAC score change (improvement) from baseline at the 2, 4, 8, 16 and 26 Week time points.

OMERACT-OARSI Responder Rate:

The OMERACT-OARSI responder rate will be derived from the WOMAC A (Pain) and C (Function) subscale scores and PGA score, as discussed in Pham (2003)[1]. Response will be determined at each visit and with the available data (no imputation of missing data will be performed to assess response). To be considered a Responder at a visit either criteria 1 or 2 must be met:

1. WOMAC A (Pain) or WOMAC C (Function) relative change from baseline ≥50% and an absolute change ≥20
2. Improvement in at least 2 of the following:
    - WOMAC A (Pain) Score: a relative change from baseline ≥20% and an absolute change from baseline ≥ 10.
    - WOMAC C (Function) Score: a relative change from baseline ≥20% and an absolute change from baseline ≥ 10.
    - Patient's global assessment (PGA): a relative change from baseline ≥20% and absolute change from baseline ≥ 10.

In 1 and 2 above, relative change will be calculated as:

$100\% \times \frac{\text{Score at Visit} - \text{Baseline Score}}{\text{Baseline Score}}$ and absolute change will be calculated as:

$\text{Score at Visit} - \text{Baseline Score}.$

Disease Activity:

Patient Global Assessment (PGA) at the 2, 4, 8, 16, and 26 Week time points.

Health Outcomes:

Physical and Mental Composite scores of the SF-12 outcomes at study defined time-points

Rescue Medicine Consumption:

Acetaminophen use from baseline through 26 Weeks.

9.4 Safety

Safety of the therapy will be assessed by comparing the incidence of adverse events in the MONOVISC vs. Saline group through 26 Weeks post-injection, including AEs that occur during injection.

10 Hypotheses

The primary endpoint analysis will be conducted to demonstrate that the mean WOMAC A1 walking pain score change from baseline for MONOVISC is significantly better than the mean WOMAC A1 walking pain score change from baseline for Saline at 26 Weeks after the first injection. The null (H_o) and alternate (H_A) hypotheses for this test of superiority are as follows:

H_o: $\mu_{MONOVISC} \geq \mu_{Saline}$

H_A: $\mu_{MONOVISC} < \mu_{Saline}$,

where $\mu_{MONOVISC}$ is the mean WOMAC A1 walking pain score change from baseline for MONOVISC at 26 Weeks, and μ_{Saline} is the mean WOMAC A1 walking pain score change from baseline for Saline at 26 Weeks.

Decision Criterion: The primary endpoint analysis will be based upon a longitudinal model as described in Section 13.4.1. The adjusted means from this longitudinal model at 26 Weeks will be compared between treatment groups. The decision will be made to reject the null hypotheses and conclude the alternative if:

1. WOMAC A1 walking pain change from baseline adjusted mean for the MONOVISC group is less than the WOMAC A1 walking pain change from baseline adjusted mean for the Saline group and,
2. The two sided p-value for the comparison of these adjusted means is less than 0.05. Or, equivalently, if the entire 2-sided 95% confidence interval for the WOMAC A1 walking pain change from baseline adjusted mean between treatment group difference (MONOVISC minus Saline), based upon the longitudinal model estimates, is less than 0.

11 Analysis Sets

A detailed accounting of all enrolled Subjects will be documented with the following analysis sets.

Intent to Treat (ITT) Analysis Set: The ITT Analysis Set consists of all subjects who are enrolled into the study. Subjects in the ITT Analysis Set but in whom treatment is not attempted with either MONOVISC or Saline will be listed along with the reason for not being treated.

Safety Analysis Set: The Safety Analysis Set consists of all randomized subjects who were enrolled into the study and in whom treatment was attempted with either MONOVISC or Saline. Subjects will be analyzed according to the treatment which was attempted.

As Treated (AT) Analysis Set for Each Follow-up Time Point: For each follow-up time point (the 2, 4, 8, 16 and 26 Week visits), the AT Analysis Set will consist of all Safety Analysis Set subjects who have a follow-up visit within the interval (see Table 2) in which WOMAC A1 walking pain data was to have been collected. Reasons for not having a WOMAC A1 walking pain result in the interval will be documented (including deaths, withdrawals of consent, and Subjects who are past due for follow-up in the interval). Subjects will be analyzed according to the treatment which was administered.

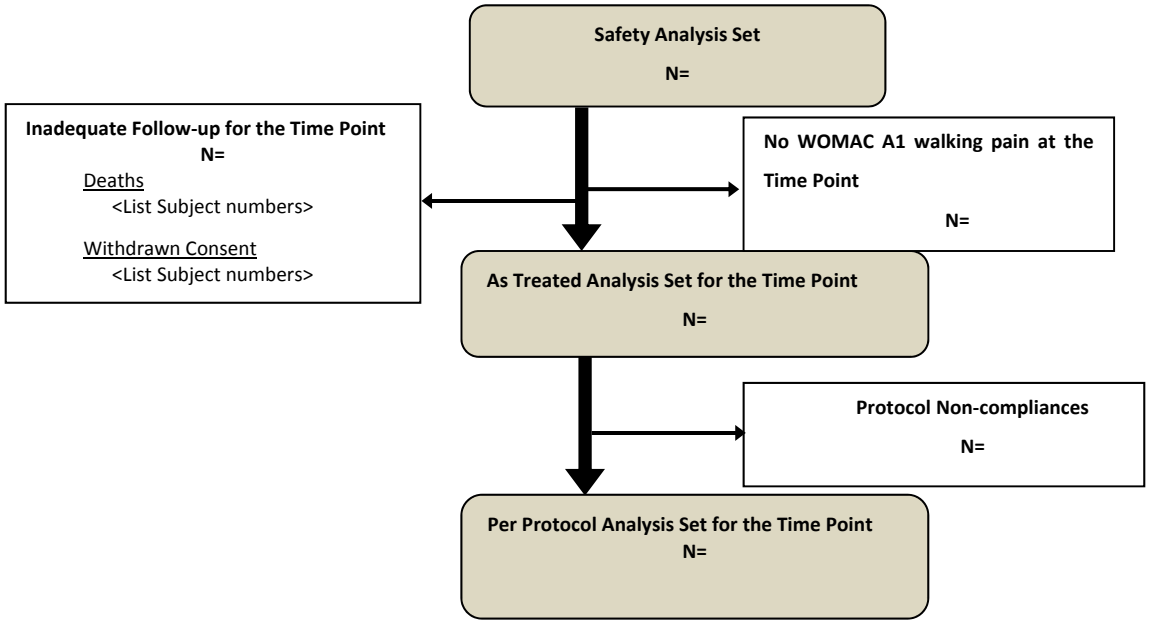
Modified Intent to Treat (MITT) Analysis Set (for Longitudinal Analysis): All primary endpoint data from all MITT Analysis Sets for Each Follow-up Time Point will be combined into a dataset for the purpose of longitudinal analysis of the primary and secondary endpoints. In this MITT Analysis Set, primary endpoint data for the 26 Week visit will be imputed with multiple imputation methodology. Subjects will be analyzed according to the treatment which was administered.

Per Protocol (PP) Analysis Set for Each Follow-up Time Point: The PP Analysis Set for each follow-up time point (the 2, 4, 8, 16 and 26 Week visits) will consist of the subjects included in AT Analysis Set for Each Time Point, excluding subjects with major protocol violations (i.e., those which are deemed to have possibly affected the scientific validity of the data). The intent is to assess inclusion of each subject at each time point independently, based upon the subject's protocol deviations. A deviation may have a transitory effect on the endpoint (e.g., use of acetaminophen within 48 hours prior to a study visit) and therefore a single time point may be impacted as opposed to study-level violations, like unmet inclusion/exclusion criterion.

Per Protocol Analysis Set for Longitudinal Analysis: All primary endpoint data from all PP Analysis Sets for Each Follow-up Time Point will be combined into a dataset for the purpose of longitudinal analysis of the primary endpoint. No PP Analysis of the Secondary endpoints will be conducted.

For each follow-up time point (the 2, 4, 8, 16 and 26 Week visits), a flowchart, as shown in Figure 2, will be created. The flowchart will present the Safety Analysis Set, the As Treated Analysis Set, and the Per Protocol Analysis Set for the Time Point. This diagram will indicate all protocol violations, deaths, withdrawals of consent, and Subjects who are past due for follow-up in the interval.

Figure 2 Example of Analysis Set Flow for a Specific Time Point



A Medically Trained Professional and a Biostatistician at DePuy Synthes will review and confirm all protocol non-compliances prior to the IA data cut, and also prior to Final dataset lock. The Analysis Sets constructed for the IA will be designated as “IA” as they will differ from the Final (end of study) Analysis Sets.

12 Sample Size Justification

The sample size was determined based on the primary endpoint, which is a superiority comparison of WOMAC A1 walking pain change from baseline at 26 Weeks for MONOVSIC Subjects compared to Saline Subjects. Although it is planned to conduct the primary endpoint analysis with a longitudinal model, and to compare the WOMAC A1 walking pain change from baseline adjusted means at 26 Weeks from this longitudinal model, the sample size was established based upon estimates for a t-test comparison of 26 week means for simplicity.

Specifically, it is anticipated that the mean difference in WOMAC A1 walking pain change from baseline means will be at least 0.6 (where the mean for MONOVISC subjects is less than the mean for Saline Subjects), and that the standard deviation in WOMAC A1 walking pain change from baseline at 26 weeks will be approximately 2.2 in both groups. This standard deviation of 2.2 is consistent with published studies by Spitzer (2010)[2] and Qvistgaard (2006)[3]; also, 2.2 is approximately ¼ of the maximum anticipated range of WOMAC A1 walking pain change from baseline, which would be an estimate of the standard deviation from data which are normally distributed. Using a two-tailed t-test and assuming a 2:1 randomization, the required sample size to yield 80% statistical power was estimated to be 318 MONOVISC and 159 Saline. With an anticipated attrition rate of 15%, the total study population was increased to be 560 patients (374 MONOVISC, 186 Saline). The sample size estimate was calculated with the following SAS code.

Figure 3 SAS Code for Power Estimation: WOCAC A Pain at 26 Weeks

```
proc power;
   twosamplemeans
   meandiff= .6
   stddev=2.2
   groupweights=(2 1)
   power= .8
   alpha=0.05
```

13 Analysis Plan

All statistical processing will be performed using SAS® Version 9.2 or higher, unless otherwise noted.

Descriptive statistics for dichotomous/categorical variables will include the number and percentage of subjects. Fisher's exact test 2-sided p-values will be provided for the comparison of dichotomous variables, including adverse event rates, to facilitate clinical judgment. Descriptive statistics for continuous variables will include the number of subjects, mean, standard deviation, median, minimum, and maximum. P-values for comparing continuous endpoints will be provided to facilitate clinical judgment; these will be 2-sided p-values based upon the t-distribution.

Summary tables will be provided for subject demographics and baseline variables for the Safety Analysis Set, the As Treated Analysis Set at 26 Weeks, and the Per Protocol Analysis Set at 26 Weeks. Safety endpoints (adverse events) will be summarized on the Safety Analysis Set. Effectiveness endpoints will be summarized at each protocol-specified evaluation time point on the Per Protocol Analysis Set for the time point. Analyses will be conducted on pooled data from all sites unless the analysis described in Section 13.7 determines sites are not poolable.

13.1 General Reporting Conventions

After all subjects have completed the primary endpoint (26 week) follow-up visit, a clinical study report will be provided to FDA for review. Trial results will also be posted on www.clinicaltrials.gov.

13.2 Subject Disposition/Patient Accountability

In addition to the Analysis Set Flowcharts as described in Section 11, the number and percentage of subjects in the following categories will be tabulated and presented by treatment group:

- Enrolled (ITT);
- Randomized;
- Safety Set;
- MITT Analysis Set for Longitudinal Analysis;
- PP Analysis Set for Longitudinal Analysis;
- MITT Analysis Set for Each Time Point;
- PP Analysis Set for Each Time Point;
- Completed Treatment;
- Completed Study Visits (2, 4, 8, 16 and 26 Week);
- Prematurely discontinued the study and the reason for discontinuation.

The disposition summary will be based on all subjects and discontinuation analysis will be based on the MITT.

13.3 Demographic and Baseline Characteristics

Summary tables will be provided for subject demographics and important baseline variables for the Safety Analysis Set, the As Treated Analysis Set at 26 weeks, and the Per Protocol Analysis Set at 26 weeks.

13.4 Efficacy Evaluations

13.4.1 Primary Endpoint Analyses

The primary endpoint analysis will be conducted with a longitudinal model on the MITT Analysis Set using PROC MIXED. The model will include absolute change from baseline in WOMAC A1 as the dependent variable, treatment (MONOVISC versus Saline) and visit (Week 2, Week 4, Week 8, Week 16, and Week 26) as fixed effects, and subject as a random effect. In the model, visit will be treated as class variable and an antidependence covariance matrix will be assumed to model the within-subject variability between visits. The antidependence structure allows the correlation between different time periods to vary. Denominator degrees of freedom will be estimated using the Kenward–Roger approximation.

In order to test the robustness of study results to the antedependence model assumptions, the primary endpoint longitudinal analysis will be repeated with two additional covariance structures: autoregressive (AR) and Toeplitz. These analyses will be conducted on the same MITT analysis set which was utilized for the primary endpoint analysis.

With a mixed effects model as the primary analysis model, no imputation of missing data at visits prior to the Week 26 will be done. However, missing data at Week 26 will be imputed, as discussed in Section 8. A supportive analysis of the primary endpoint analysis will be conducted on the Per Protocol Analysis Set for Longitudinal Analysis with the same longitudinal model.

Age, Gender, and pre-injection WOMAC A1 pain score will be compared between treatment groups and if they are found to be significantly different ($p<0.05$) in the As Treated Analysis Set at 26 weeks, then an additional sensitivity analysis will be conducted for the primary endpoint analysis. In this analysis, fixed effects for Age, Gender, and the continuous baseline value of WOMAC A1 will be added, but otherwise the model will remain the same.

If marked differences are observed between the results obtained with the primary longitudinal model and any of the sensitivity analysis, they will be reported and addressed in the Clinical Study Report.

13.4.2 Secondary Endpoint Analyses

The secondary endpoints are the WOMAC A1 walking pain change from baseline at each post-baseline time point: 2, 4, 8, and 16 Week. Each of the secondary endpoints will be analyzed utilizing the primary longitudinal model described in the Primary Endpoint Analysis (Section 13.4.1). A gate-keeping procedure will be implemented to control the family-wise Type 1 Error (see Section 6.1), thus the secondary endpoints will be tested in sequence only provided as a statistically significant result was obtained in the previous test.

If any of the sensitivity analyses conducted on the primary endpoint produces results which differ substantially from the primary analysis, the sensitivity analyses will be repeated on all secondary endpoints.

13.4.3 Additional Efficacy Analyses

No additional efficacy analysis of the primary and secondary endpoints is planned; tertiary endpoint analyses are described in Section 13.8.

13.5 Safety Evaluations

The safety of the therapy will be assessed by comparing the incidence of adverse events in the MONOVISC group vs the Saline group through 26 weeks post-injection, including AEs that occur during injection. In addition, the number and percentage of subjects with adverse events will be reported.

13.5.1 Adverse Events

An Adverse Event (AE) is any untoward medical occurrence in a subject, regardless if there is a relationship between the adverse events and the study devices.

At each evaluation of the subject enrolled in a clinical investigation, the Investigator determines whether any A) have occurred, and determines their relationship to the study devices or procedure.

All adverse events, study device malfunctions and other product issues must be recorded in the medical records and entered into CRFs. There are immediate post-injection or peri-injection events that are changes from the baseline condition of the Subject, but are expected events resulting from the treatment. If these events occur, they should be recorded in the Subject's medical record.

An Unanticipated Adverse Device Effect (UADE) is defined as any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with, a device, if that effect, problem, or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application (including a supplementary plan or application), or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects.

Serious Adverse Events (SAEs) are defined as any AE that:

1) Led to a death;

2) Led to a serious deterioration in the health of the subject that:

- Resulted in life-threatening illness or injury,
- Resulted in permanent impairment of a body structure or a body function,
- Required hospitalization or prolongation of existing hospitalization,
- Resulted in medical or surgical intervention to prevent permanent impairment to body structure or a body function;

3) Led to fetal distress, fetal death or a congenital abnormality or birth defect.

Note: A planned hospitalization and/or medical intervention for pre-existing conditions, or a procedure required by the protocol, without serious deterioration in health, is not considered to be a serious adverse event.

As MONOVISC is an approved device for the treatment of OA in the knee, certain AE are known to be associated with single intra-articular knee injections of MONOVISC. These can be found in the MONOVISC Package Insert[4]. These events are expected to be similar for hip injections; however, risks may be elevated in the present study since the protocol calls for two injections of MONOVISC rather than one injection. In the clinical trial for MONOVISC, adverse events that were related to the injection treatment were the following and are considered Anticipated AE:

- injection site pain/swelling;
- joint stiffness / swelling / effusion;
- arthralgia, or aggravated osteoarthritis;
- pain in extremity;
- synovitis;

- contusion;
- subcutaneous nodule;
- Baker's cyst.

AE will be summarized as followings with the number of events and the number and proportion of subjects with events. This analysis will be conducted using the Safety Analysis Set:

- All AE;
- All Device or Procedure Related AE (defined as possible, probable, or related);
- All AE leading to treatment discontinuation;
- All AE by Severity;
- All AE by Relationship (to investigational products or any protocol mandated procedures);
- All SAE;
- All UADE;
- All Anticipated AE.

13.5.2 Radiographs

Radiographs are not collected post-baseline and will therefore not be summarized.

13.5.3 Study Specific Data

No additional study specific data will be analyzed.

13.6 Site Heterogeneity

For the primary endpoint data at 26 Weeks, a 2-way ANOVA test with interaction term will be conducted to confirm the homogeneity of 26 week data across sites.

13.7 Poolability Analysis

To assess the poolability of outcomes across sites, a sensitivity analysis to assess the homogeneity of outcomes across sites will be conducted on the primary endpoint. The analysis is identical to the primary analysis of the primary endpoint (a longitudinal model of WOMAC A1 pain change from baseline over time) with the exception of the inclusion of 2 additional fixed effects: Site and Treatment by Site interaction. Sites with fewer than 10 subjects will be pooled into a single small site grouping while others will remain unique in the model. If the overall site effect is determined to be significant, a stratified analysis will replace the primary analysis of the primary and secondary endpoints. All analyses specified for these endpoints will be conducted by site.

13.8 Additional Analyses

The tertiary endpoints will be analyzed as described below:

Walking Pain:

WOMAC A1 score mean change over 26 Weeks will be analyzed using the same longitudinal model and analysis sets as the primary endpoint. The mean changes in WOMAC A1 over 26 Weeks will be reported as the least squares means for treatment and the between treatment group difference will be estimated.

WOMAC A1 score at the 2, 4, 8, 16 and 26 Week time points will be analyzed with a similar longitudinal model to that of the primary endpoint; however, the dependent variable will become the observed value (rather than change).

WOMAC A1 score change from baseline area under the curve (AUC) will be analyzed using the trapezoidal rule to calculate the AUC for each subject. Data will be reduced to a single AUC result per subject then an appropriate between-treatment 2 sample

t-test will be conducted. If the AUC appear non-normal, then an additional t-test comparison will be done on log-transformed data.

Pain:

WOMAC A score at the 2, 4, 8, 16 and 26 Week time points will be analyzed with a similar longitudinal model to that of the primary endpoint; however, the dependent variable will become the observed value (rather than change) for the entire pain subscale.

WOMAC A score change (improvement) from baseline at the 2, 4, 8, 16 and 26 Week time points will be analyzed with a similar longitudinal model to that of the primary endpoint; however, the dependent variable will become the change from baseline for the entire pain subscale.

Pain, Stiffness and Physical Functioning:

Summary statistics for the total WOMAC score at the 2, 4, 8, 16 and 26 Week time points will provided.

Summary statistics for the total WOMAC score change (improvement) from baseline at the 2, 4, 8, 16 and 26 Week time points will be provided.

Disease Activity:

Patient Global Assessment (PGA) at the 2, 4, 8, 16, and 26 Week time points will be summarized as tallies in each ordinal category (0 through 10).

Health Outcomes:

Summary statistics for the Physical and Mental Composite scores of the SF-12 outcomes at study defined time-points will be provided.

Rescue Medicine Consumption:

Acetaminophen use from baseline through 26 Week will be analyzed by presenting the incidence of use by study time point and overall. Difference between treatment groups will be assessed using fisher's exact test. In addition, an analysis will be conducted to provide a summary of the dose (total over the entire study period and in the interval prior to the visit).

OMERACT-OARSI Responder Rate:

The response rate will be summarized at each study time point with counts and proportions.

14 Plans for Interim Analysis (IA)

14.1 The Objectives of the IA

The objective of the IA is to determine whether the study should be stopped early due to futility. The IA does not involve the possibility of stopping early for success; since there is no possibility of stopping early for success, there is no possibility of committing a Type 1 statistical error at the interim analysis, and hence no inflation of the Type 1 statistical error will occur at the conclusion of the study. The final primary endpoint analysis will be conducted with a 2-sided p-value threshold of 0.05 for determining study success (no penalty for having conducted the interim analysis with futility stopping rule).

14.2 Efficacy and Safety Variables

The efficacy variable of interest for the IA is the WOMAC A1 pain score change from baseline difference between treatment groups (MONOVISC minus Saline), expressed as an improvement.

The assessment of Safety is not an explicitly stated objective, however, sufficient safety data is required to inform a decision concerning membership in the IA PP Analysis Set at Week 26.

14.3 Timing and Frequency of the IAs

A single IA of the primary endpoint will be conducted after 74 subjects in the MONOVISC arm and 37 patients in the saline arm have entered the IA PP Analysis Set at 26 Weeks.

14.4 Analysis Sets for the IA

Two analysis sets will be utilized in the IA, the IA PP Analysis Set at 26 Weeks and the IA AT Analysis Set at 26 Weeks.

14.5 Data Cut Criteria

The IA results will inform decision-making regarding study continuation. Therefore, all data required to conduct the analysis must be clean. As the study will be ongoing at the time of the IA and data will continue accrue, a mechanism to ensure only clean data are analyzed is required. A filtering algorithm will be implemented to restrict the data analyzed to include only data reported for subjects included in union of the IA AT Analysis Set at 26 Weeks and the IA PP Analysis Set at 26 Weeks. In addition to the WOMAC data, all data required to make the determination of inclusion into the IA PP Analysis Set at 26 Weeks is required.

14.6 Analysis Plan

The analysis will consist of an assessment of the 26 Week improvement of the MONOVISC group over the Saline group in mean WOMAC A1 walking pain score. This analysis will be conducted on the IA PP Analysis Set at 26 weeks and also the IA AT Analysis Set at 26 Weeks. There is no reliance on statistical significance (p-values or confidence intervals) as the comparison is made to a predefined futility boundary.

Analysis will be conducted by a team of qualified employees from within DePuy Synthes Mitek Sports Medicine BSDM department who will be designated to conduct the IA. Members of this team will not be otherwise associated with the conduct of the study and the analysis will be conducted in a separate area with access granted only to those designated as unblinded.

14.7 Possible Decisions or Actions Based on the IA Results

An early stopping rule for futility will be implemented in which the study will be terminated if the WOMAC A1 pain score change from baseline difference (MONOVISC minus Saline) is less than 0.4 in both the IA PP Analysis Set at 26 Weeks and the IA AT Analysis Set at 26 Weeks. All Patients already enrolled in the study at the time of the interim analysis will be followed to 26 weeks. No additional enrollment will be conducted. If this difference is greater than 0.4 (which favors MONOVISC) in either the IA PPP Analysis Set at 26 Weeks or the IA AT Analysis Set at 26 Weeks, the study will continue to completion.

14.8 Statistical Approach for Modifying the Study Design

Not Applicable

15 List of Tables, Figures, and Subject Data Listings

The document which contains the planned presentation shells (tables, figures and listings) as well as further detail on scoring algorithms for the patient reported outcomes will be created prior to final database lock, and will be called
15-MVH-01 MONOVISC HIP Shells v1.0.

16 REFERENCES

[1] Pham T, Heijde D, Lassere M, et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. J Rheumatol, 30:1648–54 (2003).

[2] Spitzer, A.I. et al, Hylan G-F 20 improves hip osteoarthritis: a prospective, randomized study, The Physician and Sports Medicine, 2(38): 35-47 (2010)

[3] Qvistgaard, E. et al, Intra-articular treatment of hip osteoarthritis: a randomized trial of hyaluronic acid, corticosteroid, and isotonic saline, Osteoarthritis & Cartilage, 14(2): 163-70 (2006)

[4] MONOVISC Package Insert