

### **16.1.9 Documentation of Statistical Methods, Study 06-UPLF-01**

The SAP and any addenda for Study 06-UPLF-01 are appended.

Modifications to the pre-defined SAP are summarized in Section 9.7 of this study report.

**Analysis Plan for Prospective  
Data Collection from the Stryker  
Biotech IDE Study of OP-1®  
Putty in Uninstrumented  
Posterolateral Fusions**

**Protocol No. 06-UPLF-001**

**June 15, 2007**

**Final**

*Prepared for*  
Stryker Biotech  
35 South Street  
Hopkinton, MA 01748

*Prepared by*  
Rong Lin, M.D., M.P.H.  
Biostatistician II

Bonnie S. Bielefeld, M.S.  
Director and Principal Biostatistician

*Reviewed by*  
Kay Larholt, Sc.D.  
Vice President and Executive Director

Biometrics  
Abt Associates Clinical Trials  
181 Spring Street  
Lexington, MA 02421



## Table of Contents

Signature Page for Analysis Plan.....	i
Table of Contents.....	ii
List of Tables.....	iv
List of Listings.....	viii
1.0 INTRODUCTION.....	1
1.1 Study Design of Protocol S01-01US.....	1
1.2 Rationale for Prospective Data Collection Protocol 06-UPLF-001.....	2
2.0 STUDY OBJECTIVES.....	4
3.0 STUDY INVESTIGATIONAL PLAN.....	5
3.1 Study Design.....	5
3.2 Selection of Study Population.....	6
3.2.1 Inclusion Criteria.....	6
3.2.2 Exclusion Criteria.....	6
3.3 Study Product.....	6
3.4 Evaluation Schedule.....	7
4.0 STATISTICAL METHODOLOGY.....	7
4.1 General Considerations.....	7
4.2 Hypothesis Testing.....	7
4.3 Sample Size.....	9
4.4 Efficacy Assessments.....	9
4.5 Safety Assessments.....	15
4.6 Multiple Imputations for Handling of Dropouts or Missing Data.....	15
4.7 Adjustments for Covariates.....	19
4.8 Multiple Comparisons/Multiplicity.....	19
4.9 Multicenter.....	19
4.10 Examination of Subgroups.....	19
4.11 Adjudication of Radiographic and CT Scan Results.....	19
5.0 STATISTICAL ANALYSES.....	20
5.1 Analysis Populations.....	20
5.1.1 Modified Intent-to-Treat Population.....	20
5.1.2 Supplemental Per Protocol Population.....	20
5.1.3 Safety Population.....	20
5.2 Patient Accountability.....	20
5.3 Non-response Bias.....	20
5.4 Demographic and Baseline Characteristics.....	20
5.5 Efficacy Analysis.....	21
5.5.1 Primary Efficacy Endpoint.....	21

5.5.2	Secondary Efficacy Endpoints.....	23
5.6	Safety Analysis.....	24
5.6.1	Adverse Events.....	25
5.6.2	Analysis of Retreatments.....	26
5.6.3	Neurological Status.....	26
5.7	Additional Analysis.....	27
5.7.1	Visual Analog Scale for Pain Assessment.....	27
5.7.2	Donor Site Pain.....	27
5.7.3	General Health Survey (SF-36).....	28
5.8	Immunology.....	28
6.0	REFERENCES.....	30

### List of Tables

<u>Table</u>	<u>Title</u>
S.1.1	Patient Disposition
S.1.2	Patient Populations
S.1.3.1	Length of 36+ Month Follow-up – Safety Population
S.1.3.2	Patient Status by Time in Study – Safety Population
S.2	Select Demographics and Success Criteria for Patients Who Participated in 06-UPLF-01 and Patients Who Did Not Participate in 06-UPLF-01 - Modified Intent-to-Treat Population
S.3	Number of Patients Who Had Missing Data and Required Imputation for Overall Success Rate - Modified Intent-to-Treat Population
S.3.1.A	Overall Success Rate – Modified Intent-to-Treat Population
S.3.1.B	Overall Success Rate – Supplemental Per Protocol Population
S.3.2	Overall Success Rate by Size of Center – Modified Intent-to-Treat Population
S.3.3	Overall Success with Test of Non-inferiority Using Equivalence Limit of 0.10 – Modified Intent-to-Treat Population
S.3.4.A	Overall Success by Gender and Age Group – Modified Intent-to-Treat Population
S.3.4.B	Overall Success Rate - by Gender and Age Group – Supplemental Per Protocol Population
S.3.5.A	Overall Success Rate With Presence of Bridging Bone – Modified Intent-to- Treat Population
S.3.5.B	Overall Success With Presence of Bridging Bone – Supplemental Per Protocol Population
S.4.A	36+ Month Overall Radiographic Success Rate – Modified Intent-to- Treat Population
S.4.B	36+ Month Overall Radiographic Success – Supplemental Per Protocol Population
S.5.A	36+ Month Overall Radiographic Success Rate with Presence of Bridging Bone – Modified Intent-to- Treat Population
S.5.B	36+ Month Overall Radiographic Success Rate with Presence of Bridging Bone – Supplemental Per Protocol Population
S.6.A	36+ Month Success Rate Based on Oswestry Disability Score – Modified Intent-to-Treat Population
S.6.B	36+ Month Success Rate Based on Oswestry Disability Score – Supplemental Per Protocol Population
S.7.A	36+ Month Success Rate Based on Absence of Retreatment – Modified Intent-to-Treat Population
S.7.B	36+ Month Success Rate Based on Absence of Retreatment – Supplemental Per Protocol Population
S.8.A	36+ Month Success Rate Based on Absence of Serious Treatment-Related Adverse Events – Modified Intent-to-Treat Population
S.8.B	36+ Month Success Rate Based on Absence of Serious Treatment-Related Adverse Events – Supplemental Per Protocol Population

**List of Tables (Cont'd)**

<u>Table</u>	<u>Title</u>
S.9.A	36+ Month Overall Neurological Success – Modified Intent-to-Treat Population
S.9.B	36+ Month Overall Neurological Success – Supplemental Per Protocol Population
S.10.1.A	Presence of Bone at Operated Level Based on 36+ Month CT – Modified Intent-to-Treat Population
S.10.1.B	Presence of Bone at Operated Level Based on 36+ Month CT – Supplemental Per Protocol Population
S.10.2.A	Presence of Bridging Bone at Operated Level Based on 36+ Month CT – Modified Intent-to-Treat Population
S.10.2.B	Presence of Bridging Bone at Operated Level Based on 36+ Month CT – Supplemental Per Protocol Population
S.10.3	Location of Bone at Operated Level Based on 36+ Month CT – Modified Intent-to-Treat Population
S.11.A	Reanalysis of Presence of Bone at Operated Level Based on 9 Month CT – Modified Intent-to-Treat Population
S.11.B	Reanalysis of Presence of Bone at Operated Level Based on 9 Month CT – Supplemental Per Protocol Population
S.12.1	Presence of Bone at Operated Level Based on 9 Month CT or 36+ Month CT – Modified Intent-to-Treat Population
S.12.2	Presence of Bone Based on 36+ Month Scout Film vs. 24-Month Film or 36+ Month CT – Modified Intent-to-Treat Population
S.13.A	Degree of Angular Motion – Modified Intent-to-Treat Population
S.13.B	Degree of Angular Motion – Supplemental Per Protocol Population
S.14.A	Translational Movement – Modified Intent-to-Treat Population
S.14.B	Translational Movement – Supplemental Per Protocol Population
S.15.A	Oswestry Disability Index - Modified Intent-to-Treat Population
S.15.B	Oswestry Disability Index – Per Protocol Population
S.16.1	Outcome and Action Taken Since Last Visit for Ongoing Adverse Events – Safety Population
S.16.2	Ongoing Adverse Events at Last IDE Visit by System Organ Class and Preferred Term – Safety Population
S.17.1	New Treatment-emergent Serious Adverse Events by System Organ Class and Preferred Term – Safety Population
S.17.2	New Treatment-emergent Serious Adverse Events by System Organ Class, Preferred Term and Severity – Safety Population
S.17.3	New Treatment-emergent Serious Adverse Events by System Organ Class, Preferred Term and Relationship – Safety Population
S.17.4	New Treatment-emergent Serious Adverse Events by System Organ Class, Preferred Term and Time Since PLF – Safety Population
S.17.5	Outcome and Action Taken For New Treatment-emergent Serious Adverse Events – Safety Population

**List of Tables (Cont'd)**

<u>Table</u>	<u>Title</u>
S.18.1	New Medical History or Physical Findings by System Organ Class and Preferred Term – Safety Population
S.18.2	New Medical History or Physical Findings by System Organ Class, Preferred Term and Severity – Safety Population
S.18.3	New Medical History or Physical Findings by System Organ Class, Preferred Term and Severity – Safety Population
S.18.4	Outcome and Action Taken For New Medical History or Physical Findings – Safety Population
S.19.1	Surgical Intervention to Treated Spinal Level – Safety Population
S.19.2	Surgical Intervention to Treated Spinal Level by Time Since PLF – Safety Population
S.20.1	Shifts in Neurological Status from Baseline to 36+ Months: Muscle Strength – Safety Population
S.20.2	Shifts in Neurological Status from Baseline to 36+ Months: Reflexes – Safety Population
S.20.3	Shifts in Neurological Status from Baseline to 36+ Months: Straight Leg Raises – Safety Population
S.20.4	Shifts in Neurological Status from Baseline to 36+ Months: Sensory Evaluation – Safety Population
S.21	Visual Analog Scale for Pain Assessment – Actual Value and Change from Baseline – Safety Population
S.22.1	Donor Site Pain – Visual Analog Scale – Safety Population
S.22.2	Donor Site Pain Status – Safety Population
S.23.1	SF-36 Health Survey Scale: Physical Component Score – Safety Population
S.23.2	SF-36 Health Survey Scale: Mental Component Score – Safety Population
S.23.3	SF-36 Health Survey Scale: Physical Functioning Scale – Actual Value and Change from Baseline – Safety Population
S.23.4	SF-36 Health Survey Scale: Role-Physical Scale – Safety Population
S.23.5	SF-36 Health Survey Scale: Bodily Pain Scale – Safety Population
S.23.6	SF-36 Health Survey Scale: Mental Health Scale – Safety Population
S.23.7	SF-36 Health Survey Scale: Role-Emotional Scale – Safety Population
S.23.8	SF-36 Health Survey Scale: Social Functioning Scale – Safety Population
S.23.9	SF-36 Health Survey Scale: Vitality Scale – Safety Population
S.23.10	SF-36 Health Survey Scale: General Health Perceptions Scale – Safety Population
S.24.1	Anti-OP-1 Antibody Status at 36+ Month Visit – Safety Population
S.24.2	Anti-OP-1 Titer – Safety Population
S.24.3	Profile for Patients with Neutralizing Antibodies – Safety Population
S.24.4	Success Outcome by Neutralizing Antibody Status – Safety Population
S.24.5	New Treatment-emergent Serious Adverse Events by Neutralizing Antibody Status and Visit – Safety Population

**List of Tables (Cont'd)**

<u>Table</u>	<u>Title</u>
S.24.6	New Immunologically-related Treatment-emergent Serious Adverse Events by Neutralizing Antibody Status and Visit – Safety Population

### List of Listings

<u>Listing</u>	<u>Title</u>
1	Patient Participation
2	Inclusion Criteria
3	Physical Exam and Medical History
4	Neurological Evaluation
5	Oswestry Low Back Pain Disability Questionnaire
6	Visual Analog Scale for Pain Assessment
7	General Health Survey (SF-36)
8	Status of Adverse Events Report as Ongoing at Last Visit of Pivotal IDE Study
9	New Adverse Events
10	Surgical Procedures Resulting from Adverse Events
11.1	Radiographic Evaluation on CT – 36+ Months
11.2	Radiographic Evaluation on CT – Reread of 9-Month Film
12	Radiographic Evaluation – Flexion and Extension Film
13	Immunology
14	Antibody Status and Adverse Events – Patients with Any Antibody Status Data
15	Patients in Pivotal IDE Study with Missing Overall Success Data After Use of 36+ Month Radiographic Data
16	Patient Efficacy

## **1.0 INTRODUCTION**

This document details the analysis plan for prospective data collection from the Stryker Biotech Pivotal IDE Study of OP-1® Putty in Uninstrumented Posterolateral Fusions (Protocol S01-01US). It describes the proposed safety and efficacy analyses, including planned summary tables and by-patient listings.

This protocol includes prospective data collection with appropriate statistical analysis on patients who have participated in the Pivotal IDE study and are currently at least three years beyond study treatment, in order to provide additional efficacy and safety data on OP-1 Putty in uninstrumented posterolateral fusion (PLF). All clinical measurements that were previously evaluated in these patients will be collected, with the exception of laboratory assessments and anterior-posterior (AP) and lateral plain films. Since it is not possible to evaluate the patients at a standardized time point, data collected after 36 months under the current protocol, will be collectively analyzed and referred to as 36+ months in all analyses.

The radiographic modality used to visualize the presence of new bone formation in the Pivotal IDE study (plain film radiography) was not sufficiently sensitive for estimating the true rate of bone formation in the OP-1 Putty treatment group. Given the greater sensitivity of CT scans in assessing presence of bone, we believe it is appropriate and valid to collect follow-up CT imaging for the determination of radiographic success following treatment with OP-1 Putty for an accurate assessment of overall patient success. In addition, the measurement of translation and angulation, dynamic markers of radiographic success, has also improved substantially since the pivotal uninstrumented trial (Protocol #S01-01US). In the pivotal trial, angulation and translation were measured by the radiographic readers by hand. Currently angulation and translation are assessed by computer programs, which systematically calculate the measurements from markers placed by the readers at the appropriate anatomical locations. This new method allows for a more objective and presumably more precise measurement of both angulation and translation. We therefore propose to prospectively collect measures of radiographic success in as many patients as possible in both the OP-1 Putty treatment and autograft control arms of the Pivotal IDE study. In this protocol, we define measures of radiographic success as presence of bone by CT scan and angulation of  $\leq 5$  degrees and translation  $\leq 3$  mm.

### **1.1 Study Design of Protocol S01-01US**

Protocol S01-01US, entitled "A Prospective, Randomized, Controlled, Multicenter, Pivotal Study of OP-1 Putty in Uninstrumented Posterolateral Fusions", was a controlled, open-label (with blinded radiographic assessment), randomized, prospective, multicenter, multinational pivotal study in which patients with single level (L3-S1) degenerative lumbar spondylolisthesis (Grade 1 or 2) and spinal stenosis underwent decompression and posterolateral spinal fusion. After signing the informed consent form, and prior to the surgical procedure, patients were randomized to treatment in a 2:1 ratio to either OP-1 Putty or a control arm, in which autogenous bone graft from the iliac crest (autograft) was used.

Patients underwent standard surgical procedures for lumbar spinal posterior decompression with concomitant posterolateral intertransverse process arthrodesis using OP-1 Putty or autograft, as

determined by randomization. Patients were evaluated postoperatively at 6 weeks and 3, 6, 9, 12, and 24 months, and annually thereafter, until the last patient achieved 2 years of follow-up.

A total of 336 patients were enrolled and randomized, and 295 were treated at 24 US and Canadian clinical sites: 208 received OP-1 Putty and 87 received autograft. The remaining 41 enrolled patients withdrew prior to study treatment.

Overall Success was defined in the Statistical Analysis Plan (SAP) for that protocol as a composite measure with the following components, determined at 24 months:

- Improvement of at least 20% in the Oswestry Disability Index (ODI) from baseline
- Absence of retreatment
- Absence of treatment-emergent serious adverse events (SAEs)
- Absence of a decrease in neurological status (assessing muscle strength, reflexes, sensory and straight leg raise), unless attributable to a concurrent medical condition or to the surgical procedure by a blinded Independent Neurological Reviewer.
- Radiographic demonstration of spinal fusion, which was also a composite measure comprising all of the following:
  - Presence of bone formation
  - Angulation of  $\leq 5^\circ$  on flexion/extension radiographs of the affected level
  - Translational movement of  $\leq 3$  mm on flexion/extension radiographs of the affected level

## 1.2 Rationale for Prospective Data Collection Protocol 06-UPLF-001

The Pivotal IDE study results showed an apparent disparity at 24 months between the positive clinical outcome reported at 24 months and the lack of radiographic success in the OP-1 Putty treatment group compared to the autograft control arm. Upon additional review of imaging data generated from the 24-month plain radiographs and 9-month computerized tomography (CT) scans in the OP-1 Putty treatment group, it became apparent that CT scans allowed the evaluation of new bone formation with more confidence and precision. Stryker believes that plain radiographs most likely underestimated the degree of new bone formation at 24 months and that 24-month CT data would have demonstrated a much higher rate of new bone formation in the OP-1 Putty treatment group.

Further, as mentioned above, the technology for assessment of angulation and translation has improved since the design of the pivotal trial. While the technology at the time of the study's design required manual measurements of translation and angulation, the new technology provides partially automated computation of these parameters and thus more objective and presumably more precise measurements.

This study will seek to expand the information regarding efficacy as well as the long-term safety of OP-1 Putty in uninstrumented posterolateral fusion. In the Pivotal IDE study, only a small number of patients in each group (34 in OP-1 Putty and 14 in autograft) were assessed clinically

at 36 months. This protocol will capture 36+ month data on all eligible patients to enable a more extensive assessment of the efficacy and patient outcomes following PLF with either OP-1 Putty or autograft.

Radiographic imaging is the most commonly used means of visualizing new bone formation, but the images can be affected by the geometry and location of the implants. The interpretation of new bone formation by plain film radiographs can be subjective and complicated by radiographic lucencies in the implant area. Computed tomography imaging is now widely accepted as the most precise and accurate technique to confirm new bone formation since CT scans eliminate overlapping and rotational errors that occur with plain-film radiography and allow three-dimensional visualization of the bone formation.<sup>1</sup> Presence of bone will be defined by new bone formation seen on axial CT Scans collected during this study and also by 3-Dimensional reconstruction views.

When the protocol was developed for the Pivotal IDE study, CT scanning was not the standard mode for collection of radiographic data for demonstrating new bone formation. Therefore, data from CT scans were not incorporated into the primary radiologic measure of success at 24 months and were only obtained at 9 months for supportive information.

After evaluation of the pivotal trial data, it became clear that there was a curious discrepancy between the positive clinical outcomes of the patients treated with OP-1 and the presence of new bone formation on these patients. Therefore, an OP-1 and autograft subset of patients heavily weighted towards radiographic presence of bone failures by plain x-ray and 9 month CT were evaluated by re-reading their 9-month CT scans for presence of bone. These re-reads were done by two blinded spine surgeons and 1 radiologist with no prior knowledge of the study. They were instructed to look for bone both medially as well as near the transverse process. This subset of the subjects clearly demonstrated that OP-1 Putty-directed new bone formation may appear along the lateral border of the superior and inferior facets, the pars interarticularis and the medial portion of the transverse processes rather than along the entire superior and inferior borders of the transverse processes as previously thought. This pattern suggests that the bone formation would not be adequately observed by simple plain-film radiographs. Thus, the radiographs collected in study S01-01US may have been relatively insensitive to detecting new bone formation.

Given the greater sensitivity of CT scans in assessing presence of bone, we believe it is imperative to collect follow-up CT images for the determination of radiographic success following treatment with OP-1 Putty for an accurate assessment of overall patient success. Therefore, this protocol will prospectively collect follow-up CT scans in as many patients as possible in both arms of the Pivotal IDE study. All patients are at least 3 years post procedure at this time, and some are out as far as 5 years.

---

<sup>1</sup> Guidelines for the Performance of Fusion Procedures for Degenerative Disc Disease of the Lumbar Spine, Journal of Neurosurgery: Spine, June 2005, volume 2, Number 6 page 636-740.

In addition, due to the enhanced precision of current angulation and translation measurement techniques as well as the fact that a significant proportion of patients with pseudoarthrosis are not detected by two years with these measurements, we propose to prospectively collect all radiographic success measures in as many patients as possible in both the OP-1 putty and autograft arms of the pivotal study.

Further, all patients enrolled in this study will be brought back for re-assessments of all clinical parameters including ODI assessments, neurological testing, and determination of surgical retreatment (i.e., revision, removal, supplemental fixation or reoperation) at the original treated level. We will also obtain serum for assessment of OP-1 immunogenicity at the follow-up visit in patients who were antibody positive at their last recorded visit, and assess the safety of OP-1 Putty by capturing follow-up from on-going adverse events (AE), occurrence of new medical conditions/physical findings, and occurrence of new serious adverse events (SAE) since the last Pivotal IDE study visit.

Collection of the same clinical and radiographic data as in the pivotal IDE study, with the exception of a helical CT scan substituted for the plain x-ray, will permit analyses of each of the components of the primary efficacy endpoint at 36+ months.

The proportion of patients with presence of new bone formation by CT scan at 9 months in the subset re-read after the completion of the original pivotal trial analysis was substantially higher than in the earlier read. This has a couple of possible explanations. First, the 9 month CT scans obtained in the pivotal IDE study were not included as part of the primary endpoint. Therefore, the CT scans were not properly adjudicated. Most importantly, the instructions for reading were not optimal, given what we now know about OP-1's medial displacement leading to new bone formation along the lateral borders of the superior and inferior facets, the pars interarticularis and the medial portion of the transverse processes rather than along the entire superior and inferior borders of the transverse processes as previously thought. Thus, in order to support our PMA package, we plan to have two blinded spine surgeons re-read the 9 month CT scans; if there is disagreement, a third reader will independently read the CT scans. Presence of bone will be determined by the majority of readers. These 9 month CT scan results will then be compared to the 36+ month CT results to allow the construction of a two-by-two table for each treatment group to demonstrate both the consistency and the durability of effect on presence of bone from 9 months to 36+ months. This information would then allow one to conclude whether bone was present by 9 months in both arms as well as whether it persisted to 36+ months in the vast majority of patients. Thirty-six plus month digitized scout films for the CT (directly comparable to 36+ month plain x-rays) will also be evaluated and compared to both 24 month plain films and 36+ month CT scan results to evaluate the potential improved sensitivity of CT scans at assessing new bone formation in this patient population.

## 2.0 STUDY OBJECTIVES

The objective of this prospective data collection is to provide additional data to support the safety and efficacy of OP-1 Putty as a replacement for autograft in patients undergoing posterolateral spinal fusion as measured by:

- Collecting additional radiographic data on all patients at least three years beyond (36+ months) initial treatment in this study. The prospectively collected data will be used to analyze primary efficacy as measured by calculating overall success rates with clinical data previously collected and analyzed but using the new 36+ months radiographic data in the OP-1 Putty and the autograft groups.
- Secondly, all data, including clinical outcome and safety data at 36+ months or more beyond initial treatment, will be used to complement the primary analysis.
- Additionally, the consistency between presence of bone at the 9-month CT and the 36+ month CT will be examined.

The principal goal of this current protocol is to show that patients treated with OP-1 Putty not only experienced clinical benefits comparable to those of patients receiving autograft, but also that these patients have comparable evidence of new bone formation and dynamic radiographic stability.

### 3.0 STUDY INVESTIGATIONAL PLAN

#### 3.1 Study Design

This study is a prospective data collection at a single visit 36+ months post-treatment of patients from the Pivotal IDE study for the purpose of obtaining additional efficacy and safety data on as many patients as possible treated in the Pivotal IDE study.

All eligible patients will be contacted for participation in the study. The maximum number of eligible patients (excludes deaths and retreatment failures) for the Pivotal IDE study is shown in the following table.

Treatment Group	Treated	Deaths*	Retreatments*	Eligible for Follow-up
OP-1 Putty	208	7	18	183
Autograft	86**	4	8	74

\*Includes all deaths or retreatments reported during the Pivotal IDE study (i.e., may include deaths/retreatments after 24 months)

\*\*Excludes patient 4116, who was treated in the Pivotal IDE study, but did not have any post-baseline visits.

### **3.2 Selection of Study Population**

#### **3.2.1 Inclusion Criteria**

To be eligible for study participation, a patient must meet all of the following criteria:

1. The patient was treated in Stryker Biotech clinical protocol S01-01US (Pivotal IDE study) and was not a retreatment failure at the time of completion of the Pivotal IDE study.
2. The patient or legal guardian is willing and able to understand, sign and date the study-specific Patient Informed Consent, which has been approved by the Institutional Review Board (IRB).
3. The patient agrees to complete the necessary clinical and radiographic evaluations. Radiographic evaluations will not be required if the patient is pregnant.

#### **3.2.2 Exclusion Criteria**

There are no exclusion criteria for participation in this protocol. Patients who died subsequent to the Pivotal IDE study will be considered missing for analyses of data at 36+ months.

### **3.3 Study Product**

The investigational product being evaluated in this study is OP-1 Putty. OP-1 Putty is composed of recombinant human osteogenic protein (rhOP-1), type I bovine bone collagen matrix, and a putty additive of carboxymethylcellulose sodium (CMC), an anionic cellulose derivative that yields a putty-like consistency. OP-1 Putty is provided as two components:

- A large vial containing a sterile dry power consisting of 3.5 mg of human recombinant osteogenic protein-1 (OP-1) in 1 gm of collagen matrix
- A small vial containing the Putty additive consisting of a sterile dry powder composed of 230 mg CMC

For the OP-1 Putty arm, one product unit was used on each side of the spine, i.e., two product units per patient.

The control system being utilized in this study was lumbar spinal fusion with the use of autogenous bone graft from the iliac crest (autograft).

### 3.4 Evaluation Schedule

The current study is designed as a single visit, but the study assessments may be completed in a single visit or multiple visits (e.g., clinical evaluations and radiographic evaluations performed on separate dates) after obtaining informed consent. The patient's participation in this study is considered complete following collection of all specified evaluations.

## 4.0 STATISTICAL METHODOLOGY

### 4.1 General Considerations

All summary tables will be produced and all statistical analyses will be performed using SAS software.

Continuous variables will be summarized using descriptive statistics, specifically the mean, median, standard deviation, minimum and maximum. Categorical variables will be summarized using frequencies and percentages.

Inferential tests will be performed at the 5% level of significance. All  $p$ -values will be rounded to 3 decimal places. If a rounded  $p$ -value is 0.000 (i.e., the actual  $p$ -value is less than 0.0005), then this will be presented as a  $p$ -value of '<0.001'.

Information displayed in the tables generally will be presented for both treatments (the OP-1 Putty group and the autograft group).

Data listings will be presented only for data collected during this current protocol and will present data for all patients, sorted by treatment and patient ID. All date fields will be presented in a format of ddmmmyyyy (e.g., 01Jan2004) in the listings. Additionally, derived outcome assessments for all patients in the Pivotal IDE study will be updated and included in the listings for this current protocol.

### 4.2 Hypothesis Testing

This study will seek to expand the information regarding efficacy as well as the long-term safety of OP-1 Putty in uninstrumented PLF. In the Pivotal IDE study, only a small number of patients in each group (34 in OP-1 Putty and 14 in autograft) were assessed clinically at 36 months. This protocol will capture 36+ month data on all eligible patients to enable a more extensive assessment of the efficacy and patient outcomes following PLF with either OP-1 Putty or autograft.

The null hypothesis for this study is the same as for the Pivotal IDE study, specifically that the difference between the success rate in the autograft treatment group ( $P_A$ ) and the success rate in the OP-1 Putty treatment group ( $P_O$ ) is greater than or equal to the non-inferiority margin  $\delta_p$  ( $P_A - P_O \geq \delta_p$ ). Likewise, the alternative hypothesis for this study is the same as for the Pivotal

IDE study, specifically that this difference is less than  $\delta_p$  ( $P_A - P_O < \delta_p$ ). If the null hypothesis of inferiority of OP-1 is rejected, a test of superiority will be performed.

The actual testing of the hypothesis will involve the “angular transformation” (a standard transformation in statistical analyses of proportions, developed to remove the dependence of the variability of observed proportions on the underlying proportion ---  $P_A$  or  $P_O$ ). Mathematically, this transformation begins with the observed proportion of successes,  $x/n$  from  $x$  successes in  $n$  trials, and applies the function  $\sin^{-1} \sqrt{x/n}$  to produce the corresponding angle (in radians). Because the statistical variation in these angles is constant across the angular scale (in sizable samples the variance is  $1/(4n)$ ), a non-inferiority margin  $\delta_A$  that is constant in the angular scale has uniform impact, relative to the variation. Further, because the distribution in the angular scale is well approximated by a normal distribution, the customary hypothesis test in that scale (based on the normal distribution) will have the specified Type I error rate, .05.

If the angular values corresponding to  $P_A$  and  $P_O$  are  $A_A$  and  $A_O$ , respectively, then the null hypothesis becomes  $A_A - A_O \geq \delta_A$ . Specifying a constant non-inferiority margin in the angular scale ( $\delta_A$ ) corresponds to allowing the non-inferiority margin in the proportion scale,  $\delta_p$ , to vary. The relation between  $\delta_p$  and  $\delta_A$  is anchored at  $P_A$  and  $A_A$ :  $A_A = \sin^{-1} \sqrt{P_A}$ , subtracting  $\delta_A$  gives  $A_O = A_A - \delta_A$ , inverting the angular transformation yields  $P_O = (\sin(A_O))^2$ , and  $\delta_p = P_A - P_O$ . For an appropriate choice of  $\delta_A$ , the corresponding values of  $\delta_p$  are lower when the underlying statistical variation (in the proportion scale) is smaller and higher when that statistical variation is larger. The testing of the null hypothesis in this study will use  $\delta_A = 0.14$ . The choice of  $\delta_A = 0.14$  maintains  $\delta_p$  close to 0.1 for the extreme values of the success rates, and is greater in the middle of the scale to allow for the greater corresponding variability. The non-inferiority margin  $\delta_p$  thus has the following values:

$P_A$	$\delta_p$
0.1	0.0673
0.2	0.0989
0.3	0.1189
0.4	0.1315
0.5	0.1382
0.6	0.1393
0.7	0.1344
0.8	0.1222
0.9	0.0985

Additionally, although not a primary or secondary endpoint, this study will test the null hypothesis that the difference between the success rate in the autograft treatment group and the success rate in the OP-1 Putty treatment group is greater than or equal to the non-inferiority margin of 10% ( $P_A - P_O \geq 10\%$ ). For this test, the alternative hypothesis will be that this

difference is less than 10% ( $P_A - P_O < 10\%$ ). This hypothesis was defined in the original protocol and is included here as per request by the FDA.

The following table summarizes which analyses will be tested for non-inferiority.

<b>Protocol 06-UPLF-001 Post-text Table (Section 14.2)</b>	<b>Title</b>
S.3.1.A	Overall Success Rate – mITT Population
S.3.3	Overall Success Rate with Test of Non-inferiority Using Equivalence Limit of 0.10 – mITT Population

#### 4.3 Sample Size

All eligible patients from the Pivotal IDE study will be contacted for participation in the current study. The maximum number of eligible patients (excludes deaths and retreatment failures subsequent to data lock for the Pivotal IDE study) will be 257 (183 and 74 for the OP-1 Putty group and autograft group, respectively).

#### 4.4 Efficacy Assessments

Efficacy will be assessed by calculating the overall success rates in the OP-1 Putty and the autograft groups. The data for the assessments of success will generally come from the Pivotal IDE study, with the exception of radiographic data (CT, angulation, and translational movement) and data on retreatments, which will be assessed during the current study.

The following assessments will be used for the patient success criteria:

- Radiographic assessment of bone by CT scan, angular motion, and translational movement at 36+ months
- Oswestry Disability Index at 24 months
- Retreatment by 36+ months, reported as revision; removal; supplemental fixation; reoperation
- Neurological evaluation at 24 months
- Absence of serious treatment-related adverse event before or at 24 months

For patients who report a retreatment to the operative spinal level prior to the 36+ month assessment, the clinical outcome assessment assigned during this current study will be set to be a failure, regardless of the classification assigned during the Pivotal IDE study, and whether or not the patient participated in the current study. This is because radiographic measures at 36+ months are meaningless if there has been a retreatment.

Patients who do not participate in this current protocol will have all 36+ month assessments set to missing, unless otherwise determined to be a failure.

Table 1 highlights changes in efficacy analyses between the Pivotal IDE study and study 06-UPLF-001.

**Table 1: Changes in Efficacy Criteria Between Pivotal IDE Study and Study 06-UPLF-001**

Efficacy Outcome	Protocol 06-UPLF.001 Post-text Table (Section 14.2)	Criteria Used During Pivotal IDE Study	Criteria for Use in Study 06-UPLF-001
Overall Success Rate	Table S.3 Table S.3.1.A Table S.3.1.B Table S.3.2 Table S.3.3 Table S.3.4.A Table S.3.4.B	Radiographic component of overall success used 24-month plain film	<ul style="list-style-type: none"> <li>• If retreatment occurs subsequent to 24-month visit, then success outcome at 24 months reset to failure, regardless of success outcome at 24-months in Pivotal IDE study</li> <li>• 36+ month radiographic success used in lieu of 24-month plain film</li> </ul>
Overall Success Rate using presence of bridging bone	Table S.3.5.A Table S.3.5.B	Radiographic component of overall success used 24-month plain film	<ul style="list-style-type: none"> <li>• If retreatment occurs subsequent to 24-month visit, then success outcome at 24 months reset to failure, regardless of success outcome at 24-months in Pivotal IDE study</li> <li>• 36+ month CT assessment of bridging bone, and 36+ month angulation and translational movement used in lieu of 24-month plain film</li> </ul>

**Table 1: Changes in Efficacy Criteria Between Pivotal IDE Study and Study 06-UPLF-001**

Efficacy Outcome	Protocol 06-UPLF-001 Post-text Table (Section 14.2)	Criteria Used During Pivotal IDE Study	Criteria for Use in Study 06-UPLF-001
36+ month overall radiographic success with presence of bone	Table S.4.A Table S.4.B	Not Assessed	<ul style="list-style-type: none"> <li>• If retreatment occurs subsequent to PFL, then overall radiographic success at 36+ months set to failure</li> <li>• 36+ month CT assessment of presence of bone and 36+ month assessment of angulation and translation based on plain film</li> <li>• Missing data imputed for the mITT population</li> </ul>
36+ month overall radiographic success with presence of bridging bone	Table S.5.A Table S.5.B	Not Assessed	<ul style="list-style-type: none"> <li>• If retreatment occurs subsequent to PFL, then overall radiographic success at 36+ months set to failure</li> <li>• 36+ month CT assessment of presence of bridging bone and 36+ month assessment of angulation and translation based on plain film</li> <li>• Missing data are not imputed</li> </ul>
36+ month success rate based on Oswestry disability score	Table S.6.A Table S.6.B	Not Assessed	If retreatment occurs subsequent to PFL, then success based upon ODI set to missing.
36+ month success rate based on absence of retreatment	Table S.7.A Table S.7.B	Not Assessed	If retreatment occurs subsequent to PFL, then success based on absence of retreatment set to failure.

**Table 1: Changes in Efficacy Criteria Between Pivotal IDE Study and Study 06-UPLF-001**

<b>Efficacy Outcome</b>	<b>Protocol 06-UPLF-001 Post-text Table (Section 14.2)</b>	<b>Criteria Used During Pivotal IDE Study</b>	<b>Criteria for Use in Study 06-UPLF-001</b>
36+ month success rate based on absence of serious treatment-related adverse events	Table S.8.A Table S.8.B	Not Assessed	Success criteria based upon absence of serious treatment-related adverse events prior to 36+ month assessment. If retreatment occurs subsequent to PFL, then success based on absence of serious treatment related adverse event set to failure.
36+ month overall neurological success	Table S.9.A Table S.9.B	Not Assessed	If retreatment occurs subsequent to PFL, then overall neurological success set to missing.
36+ month assessment of presence of bone at operated level	Table S.10.1.A Table S.10.1.B	Not Assessed	If retreatment occurs subsequent to PFL, then patient not assessed for presence of bone.
36+ month assessment of presence of bridging bone at operated level	Table S.10.2.A Table S.10.2.B	Not Assessed	If retreatment occurs subsequent to PFL, then patient not assessed for presence of bridging bone.
Reanalysis of 9 month CT reported in Protocol S01-01US	Table S.11.A Table S.11.B	No Assessed	<ul style="list-style-type: none"> <li>• CT images to be re-read.</li> <li>• Patients with retreatment prior to the 9-month CT are not included in analysis</li> </ul>
Presence of bone on 9 month and 36+ month CT	S.12.1	Not Assessed	<ul style="list-style-type: none"> <li>• For 9-month CT, use reanalyzed CT</li> <li>• If retreatment occurs subsequent to PFL, then patient not included in analysis</li> </ul>
Presence of bone on 36+ month digitized scout film and 24-month film or 36+ month CT scan	S.12.2	Not Assessed	If retreatment occurs subsequent to PFL, then patient not assessed for presence of bone.

**Table 1: Changes in Efficacy Criteria Between Pivotal IDE Study and Study 06-UPLF-001**

<b>Efficacy Outcome</b>	<b>Protocol 06-UPLF-001 Post-text Table (Section 14.2)</b>	<b>Criteria Used During Pivotal IDE Study</b>	<b>Criteria for Use in Study 06-UPLF-001</b>
Change from baseline in degree of angular motion at 36+ months	Table S.13.A Table S.13.B	Not Assessed	If retreatment occurs subsequent to PFL, then patient not included in analysis.
Change in translational movement at 36+ Months	Table S.14.A Table S.14.B	Not Assessed	If Retreatment occurs subsequent to PFL, then patient not included in analysis.
Change in ODI at 36+ months	Table S.15.A Table S.15.B	Not Assessed	If Retreatment occurs subsequent to PFL, then patient not included in analysis.

#### 4.5 Safety Assessments

Safety will be assessed by adverse events and neurological status. Regarding adverse events, these assessments will include updates on any adverse events that were reported as “ongoing” at the time of the last scheduled study visit under the original protocol; any serious adverse events, medical conditions or physical findings that have occurred since the last completed Pivotal IDE study visit; and occurrence of any retreatments to the original treated spinal level.

All study data (excluding personal identifiers) will be reported in data listings.

#### 4.6 Multiple Imputations for Handling of Dropouts or Missing Data

For the analysis of the primary endpoint for patient success and the analysis of overall radiographic success (overall success criterion #1), missing values will be imputed (in each treatment arm separately) using the data from the Pivotal IDE study as well as the same multiple imputation (MI) techniques as described in the Pivotal IDE study. As shown in Section 4.4, the evidence of overall radiographic success will come from the CT scan at 36+ months and plain film at 36+ months, rather than from plain film at 24 months. All other data elements not affected by retreatment (discussed below) remain the same. Patients who are considered as no longer participating in the study at a given visit, patients who have missed a visit, and patients who have missing data at a visit will be considered to have missing data at that visit.

An initial step will determine which patients require imputation. First, as mentioned in Section 4.4, patients who report a retreatment to the operative spinal level at any time up through the 36+ month visit will be considered failures on the primary efficacy endpoint. Second, any patient who is classified as a failure on a non-missing component of the primary efficacy endpoint will necessarily be a failure on that endpoint. (For this purpose, presence of bone at 36+ months and angular motion and translational movement at 36+ months are considered two separate components.) Third, any patient who is not in the preceding two categories and who has no missing components of the primary efficacy endpoint must be a success on that endpoint. Thus, the remaining patients, who have at least one missing component and success on all non-missing components, will require imputation. Overall radiographic success at 36+ months will be imputed separately from the other four components (listed in Section 4.4), which will be treated as a unit for purposes of imputation. Some patients may have missing data on one of the two components of overall radiographic success at 36+ months (presence of bone at 36+ months and angulation and translational movement at 36+ months) but success on the other component. If the total number of such patients in the two treatment arms exceeds 10, the missing component will be imputed (as described below) instead of imputing overall radiographic success at 36+ months as a unit.

The imputations of overall radiographic success at 36+ months will be based on variables selected from among the following potentially relevant covariates or predictors:

Model 1, Patients with 24-month data:

- Presence of bone on the 9-month CT scan

List C, with  $m = 24$ :

- Patient success on the other four components at  $m$  months
- Neurologic success at  $m$  months
- Oswestry Disability Index at  $m$  months
- Workers Compensation status at baseline
- At least 2 morbidities at baseline (assessed by medical history)

Additional models for imputation of overall radiographic success at 36+ months separate into two groups, according to whether patients are missing the 9-month CT scan.

Model 2a, Patients with 9-month CT scan and 36-month data, but without 24-month data:

- Presence of bone on 9-month CT scan

List C, with  $m = 36$ .

Model 2b, Patients with 9-month CT scan and 12-month data, but without 36-month data or 24-month data:

- Presence of bone on 9-month CT scan

List C, with  $m = 12$ .

Model 2c, Patients with 9-month CT scan and 36+ month data, but without 12-month data or 36-month data or 24-month data:

- Presence of bone on 9-month CT scan

List C, with  $m = 36+$ .

Model 2d, Patients with 9-month CT scan and 6-month data, but without 36+ month data or 12-month data or 36-month data or 24-month data:

- Presence of bone on 9-month CT scan

List C, with  $m = 6$ .

Model 2e, Patients with 9-month CT scan, but without 6-month or 36+ month data or 12-month data or 36-month data or 24-month data:

- Presence of bone on 9-month CT scan

Model 3a, Patients with 24-month data but without 9-month CT scan: List C with  $m = 24$ .

Model 3b, Patients with 36-month data but without 9-month CT scan or 24-month data: List C with  $m = 36$ .

Model 3c, Patients with 12-month data, but without 36-month data or 9-month CT scan or 24-month data: List C with  $m = 12$ .

Model 3d, Patients with 36+ month data, but without 12-month data or 36-month data or 9-month CT scan or 24-month data: List C with  $m = 36+$ .

Model 3e, Patients with 6-month data, but without 36+ month data or 12-month data or 36-month data or 9-month CT scan or 24-month data: List C with  $m = 6$ .

Model 3f, Patients without 6-month data or 36+ month data or 12-month data or 36-month data or 9-month CT scan or 24-month data: the proportion of patients (in the particular treatment arm) who have overall radiographic success at 36+ months, among those patients who have data from 36+ month CT scan and 36+ month plain film.

The imputations of the missing data on the other four components (as a unit) will be based on variables selected from among the following potentially relevant covariates or predictors:

Model 4a, Patients with 36-month data:

- Presence of bone on 36+ month CT scan
- Success of angulation of  $\leq 5^\circ$  at 36+ months
- Success of translational movement of  $\leq 3$  mm at 36+ months

List C with  $m = 36$ .

Model 4b, Patients with 12-month data, but without 36-month data:

- Presence of bone on 36+ month CT scan
- Success of angulation of  $\leq 5^\circ$  at 36+ months
- Success of translational movement of  $\leq 3$  mm at 36+ months

List C with  $m = 12$ .

Model 4c, Patients with 36+ month data, but not 12-month or 36-month data:

- Presence of bone on 36+ month CT scan
- Success of angulation of  $\leq 5^\circ$  at 36+ months
- Success of translational movement of  $\leq 3$  mm at 36+ months

List C with  $m = 36+$ .

Model 4d, Patients with 6-month data, but not 36+ month or 12-month or 36-month data:

- Presence of bone on 36+ month CT scan
- Success of angulation of  $\leq 5^\circ$  at 36+ months
- Success of translational movement of  $\leq 3$  mm at 36+ months

List C with  $m = 6$ .

Model 4e, Patients without 6-month or 36+ month or 12-month or 36-month data:

- Presence of bone on 36+ month CT scan
- Success of angulation of  $\leq 5^\circ$  at 36+ months
- Success of translational movement of  $\leq 3$  mm at 36+ months

If, as mentioned above, more than 10 patients have missing data on one of the two components of overall radiographic success at 36+ months and success on the other component, each such patient will have the missing component imputed (instead of overall radiographic success at 36+ months as a unit). The model will be based on the non-missing component and the predictors in whichever of Models 1 through 3e corresponds to the patient's available data. If the available data correspond to Model 3f, the proportion will be based on the non-missing component instead of overall radiographic success.

To carry out the MI for overall radiographic success at 36+ months and the MI for the other four components (as a unit), we will use a parametric model unless the data indicate that no model is required (e.g., among patients who have 12-month data and 24-month data, all successes at 12 months are also successes at 24 months). The parametric model for overall radiographic success at 36+ months will be a logistic regression where the outcome is overall radiographic success at 36+ months and the predictors are as listed under Models 1 through 3e. Similarly, to carry out the MI for the other four components (as a unit), we will use a logistic regression where the outcome is success/failure at 24 months and the predictors are as listed under Models 4a through 4e. Imputation will be done separately for each treatment arm.

The MI procedure begins by fitting the logistic regression model to data on the complete cases and estimating the parameters associated with each covariate (predictor) as well as the variance-covariance matrix. Separate models will be developed for patients in the subsets associated with each of the models listed above. Selection of the best logistic regression model will begin with identifying candidate models using a step-up approach including candidate covariates from the list described above. The choice of the best of the candidate models will be based on the Schwarz criterion (SC), which tends to produce parsimonious models.

The procedure described below will apply separately to imputation of overall radiographic success at 36+ months and imputation of success on the other four components. The results will then be combined to produce an imputed value of the primary efficacy measure.

Once the model has been selected, imputation begins by sampling a value for the vector of model parameters from a multivariate normal distribution whose mean and variance-covariance matrix equal those estimated from the complete data. Each missing observation is imputed by sampling from the conditional distribution of this observation given the sampled parameter values and the observed covariates. Here the probability of a success for the missing observation is assumed to be Bernoulli-distributed, with the Bernoulli parameter obtained by the logistic regression equation with the sampled parameter estimates. The number of imputations will be 4, reflecting Rubin's assertion that for moderate fractions of missing information (<30%), a small number of imputations (3 or 4) results in nearly fully efficient estimates of the population quantity of interest.

Within each treatment arm each of the multiple imputations will produce a completed-data estimate of the quantity to be used in the hypothesis test, along with an associated estimate of its within-imputation variance. The inferences will be based on the average of those completed-data estimates and on a variance estimate that incorporates the average within-estimate variance and the between-imputation variance. If the multiply imputed data are used in an adjustment for covariates (Sections 4.7 and 5.5.1), the model will involve data from both arms and will yield estimates of the adjusted treatment difference and associated estimates of its within-imputation variance.

For other efficacy analyses, missing values will not be imputed except as specified. Missing data will not be imputed for the safety data.

The following table summarizes which analyses will impute for missing data.

<b>Protocol 06-UPLF-001 Post-text Table (Section 14.2)</b>	<b>Title</b>
S.3.1.A	Overall Success – mITT Population
S.3.3	Overall Success Rate with Test of Non-inferiority Using Equivalence Limit of 0.10 – mITT Population
S.4	36+ Month Overall Radiographic Success – mITT Population

#### **4.7 Adjustments for Covariates**

An analysis will adjust for covariates (if a statistically significant adjustment is found) as described in Section 5.5.1.

#### **4.8 Multiple Comparisons/Multiplicity**

No multiplicity adjustment will be made. The statistical hypothesis test for the primary efficacy endpoint that was performed previously using plain film to assess presence of bone and angulation and translation (all at 24 months) is simply being revised to use what are considered to be more accurate and sensitive measurements of presence of bone using CT scans and more objective and precise measurements of angulation and translation (all at 36+ months). Thus, use of a multiplicity adjustment is not considered necessary.

#### **4.9 Multicenter**

For the primary efficacy endpoint, an analysis will be conducted to test for treatment by center interaction. If the interaction is significant, results will be presented by center.

#### **4.10 Examination of Subgroups**

Analysis of the primary efficacy endpoint, overall success at 24 months, will be presented by sex and age category (<45 years, 45-65 years, >65 years), as well as overall.

#### **4.11 Adjudication of Radiographic and CT Scan Results**

Thirty-six plus month and 9 month CT scans as well and 36+ month angulation and translation success will be reviewed and scored by two blinded spine surgeons who have had no other involvement in the study and no prior knowledge of study results. The CT scans will be read by two blinded spine surgeons and, if there is disagreement, a third reader will independently read the CT scans. Presence of bone at 9 months and 36+ months will be determined by the majority of readers. Angulation and translation success will also be determined by the two readers; if there is disagreement, then a third reader will make an independent assessment. Ultimately angulation and translation success will be determined by the majority.

## **5.0 STATISTICAL ANALYSES**

### **5.1 Analysis Populations**

#### **5.1.1 Modified Intent-to-Treat Population**

The modified intent-to-treat (mITT) population includes all patients who were randomized to the Pivotal IDE study and had at least one post-treatment visit during that study. All efficacy analyses will be conducted on the mITT population.

#### **5.1.2 Supplemental Per Protocol Population**

The supplemental per protocol population includes all OP-1 Putty or autograft treated patients who were in the per protocol population of the Pivotal IDE study and who had radiographic success measured at 36+ months. Analysis of overall patient success will be repeated on the supplemental per protocol population using descriptive statistics to aid in the interpretation of the primary efficacy analysis of the mITT population. All other efficacy analyses will also be repeated on the supplemental per protocol population, unless indicated otherwise.

#### **5.1.3 Safety Population**

The safety population includes all patients who were treated using either OP-1 Putty or autograft in the Pivotal IDE study. The safety analyses will be based on the safety population.

### **5.2 Patient Accountability**

Table S.1.1 will tabulate the number and percent of patients who were enrolled in the Pivotal IDE study, eligibility status and disposition for Study 06-UPLF-001, and S01-01US status of patients eligible for 06-UPLF-001. A summary of each patient population will be presented in Table S.1.2 by treatment group. Table S.1.3.1 will summarize length of 36+ month follow-up, while Table S.1.3.2 will tabulate patient status by time in study for patients treated in S01-01US, and patients enrolled in Study 06-UPLF-01.

### **5.3 Non-response Bias**

Select outcome criteria (presence of bone on 9 month CT, and 24-month overall success during Pivotal IDE study) will be presented in Table S.2, by treatment group, stratified by participation status in Study 06-UPLF-01.

### **5.4 Demographic and Baseline Characteristics**

Refer to CSR for Protocol S01-01US for summary statistics regarding demographics and baseline characteristics.

## 5.5 Efficacy Analysis

### 5.5.1 Primary Efficacy Endpoint

The primary efficacy endpoint of this study is the overall success rate for the mITT population using radiographic success at 36+ months in lieu of the 24-month plain x-ray. Most values are from the Pivotal IDE study, with the exception of the 36+ month radiographic data and retreatment data, which are from the current study.

A patient will be considered an overall success if all five of the following components, which were described in greater detail in the Pivotal IDE study, are met:

- 1) Radiographic success defined as meeting all three of the following conditions:
  - Presence of bone formation on CT at 36+ months
  - Angulation of  $\leq 5^\circ$  at 36+ months via plain film
  - Translational movement of  $\leq 3$  mm at 36+ months assessed by plain film
- 2) Oswestry Disability Index improvement of at least 20% from the pre-treatment visit at 24 months
- 3) No revisions, removals or supplemental fixations by 36+ months. All reoperations that were intended to promote fusion at the treated level are considered failures.

The term retreatment refers to a revision, removal, supplemental fixation, or reoperation intended to promote fusion at the treated level. Any patient who experienced a retreatment is considered a failure, regardless of the timing of the procedure. Patients who were a success at 24 months based upon the Pivotal IDE results, but who subsequently experienced a retreatment, will have their 24-month assessment reclassified as failure.

- 4) The absence of serious treatment-related adverse events before or at 24 months.
- 5) Patient was considered an overall neurological success at 24 months in the absence of a decrease in neurological status, unless attributable to a concurrent medical condition or to the surgical procedure.

The null hypothesis is that the difference between the success rate for the autograft treatment group ( $P_A$ ) and the success rate for the OP-1 Putty treatment group ( $P_O$ ) is greater than or equal to the non-inferiority margin  $\delta_p$  ( $P_A - P_O \geq \delta_p$ ). This hypothesis will be examined by estimating the difference between the success rates in the two treatment groups (expressed in the angular scale as  $A_A - A_O$ ) and the associated standard error. An upper 95% confidence bound on  $A_A - A_O$  will also be computed. If the null hypothesis of inferiority of OP-1 is rejected, a test of superiority of OP-1 will be performed. If an adjustment for covariates is made (as discussed

below), additional test(s) and confidence bound will be based on the adjusted treatment difference.

The primary analysis of the primary efficacy endpoint will be based on the mITT population, with the variable non-inferiority margin of this endpoint also being considered primary. The statistical procedure that accommodates missing data is described in detail in Section 4.6. This endpoint will also be analyzed with descriptive statistics based on the supplemental per protocol population to aid in the interpretation of the primary analysis of this endpoint.

The number and percentage of mITT patients in each treatment group with missing data and requiring imputation for overall success rate will be presented in Table S.3 to assess the potential impact of missing data for the mITT population. Fisher's exact test will be used to test the difference in number of patients with missing data between treatment groups. The patients with missing data and requiring imputation will also be listed in the data listings.

Tables S.3.1.A and S.3.1.B will summarize the overall success rate for the mITT and supplemental per protocol populations, respectively. The main analysis, on the mITT population (Table S.3.1.A), will use a logistic regression model to take into account the combined effect on success rate of baseline characteristics that show statistically significant differences (at the 0.10 significance level). The model will yield adjusted success rates (which may be more precise than the unadjusted rates). The characteristics to be considered for this analysis are:

- Age: <45 years old, 45-65 years old, >65 years old
- Clinical site
- Gender: male, female
- Level fused: L3-L4, L4-L5, L5-S1
- Grade of spondylolisthesis: Grade I or Grade II
- Prior treatment: surgical (laminectomy, facetectomy, foraminotomy, discectomy), not surgical (includes no previous treatment)
- Concurrent medical condition: metabolic bone disease and/or osteoporosis (yes/no)
- Concurrent medical condition: diabetes (yes/no)
- Workers Compensation status: no or yes (includes current, pending, litigation, and other)
- BMI (continuous variable)
- Oswestry Disability Index (continuous variable)

Inferences will be based on estimates of treatment effects adjusted for covariates and on variance estimates obtained through multiple imputation as described in Section 4.6.

Additionally, if significant, analysis of treatment by center interaction for the primary efficacy endpoint will be presented in Table S.3.1.A using logistic regression. If the interaction is significant at the 0.05 level, success rates for each treatment group will be presented by center, without pooling. Even if there is a significant interaction, the significance of the unadjusted difference between groups on the primary endpoint will be based on the pooled data.

Table S.3.2 will present overall success stratified by center (with large centers defined as  $n \geq 12$ , and small centers defined as  $n < 12$ ) for the mITT population. This stratification will be based on enrollment in study S01-01US. Missing data will not be imputed.

If there are statistically significant differences in baseline characteristics (Oswestry score, level fused, degree of angular motion, and translational movement and Workers Compensation Status) between treatment groups, the overall success rate for the primary efficacy endpoint will also be stratified by those characteristics which are statistically significant. The results will be presented in additional analysis tables. Missing data will not be imputed.

Table S.3.3 will summarize overall success using an equivalence limit of 0.10 to test for non-inferiority for the mITT population.

The overall success rate for the primary efficacy endpoint will also be presented separately for male and female and for each age category (<45 years old, 45 – 65 years old, >65 years old) in Tables S.3.4.A and S.3.4.B for the mITT and supplemental per protocol populations, respectively.

### 5.5.2 Secondary Efficacy Endpoints

The secondary efficacy endpoints for the study are overall success with bridging bone via CT at 36+ months, and translational movement and angulation at 36+ months via plain film (Tables S.3.5.A and S.3.5.B).

Additionally, although not specified in the protocol for Study 06-UPLF-001, for consistency with analyses performed for the Pivotal IDE study, the following secondary efficacy endpoints are being defined.

- 36+ month overall radiographic success (Tables S.4.A and S.4.B). Only the analyses based on the mITT population will impute for missing data.
- 36+ month overall radiographic success with presence of bridging bone (Tables S.5.A and S.5.B)
- 36+ month success rate based on Oswestry disability score (Tables S.6.A and S.6.B)
- 36+ month success rate based on absence of retreatment (Tables S.7.A and S.7.B)
- 36+ month success rate based on absence of serious treatment-related adverse events (Tables S.8.A and S.8.B)
- 36+ month overall neurological success rate (Tables S.9.A and S.9.B)
- Presence of bone at operated level based on the CT scans at 36+ months (Tables S.10.1.A and S.10.1.B)

- Presence of bridging bone at operated level based on the CT scans at 36+ months (Tables S.10.2.A and S.10.2.B)
- Location of bone at operated level based on the CT scans at 36+ months (Table S.10.3)
- Change from baseline in degree of angular motion at 36+ months (Tables S.13.A and S.13.B)
- Change from baseline in translational movement at 36+ months (Tables S.14.A and S.14.1B)
- Change from baseline in Oswestry Disability Index at 36+ months (Tables S.15.A and S.15.B)

Actual value and change from baseline to 36+ months for degree of angular motion, translational movement and Oswestry Disability Index will be summarized by treatment group. Two-sample t-tests will be used to test the difference in change from baseline between the OP-1 Putty group and the autograft group. The changes from baseline will also be examined using one-sample t-test to test the mean change within each treatment group. Baseline values will be from the Pivotal IDE study.

For degree of angular motion and translational movement at 36+ months, the median of the scores from all reviewers will be used.

Additionally, a reanalysis of the 9 month CT data collected during the Pivotal IDE study will be performed. The results will be presented based on presence of bone in Tables S.11.A and S.11.B for the mITT and supplemental per protocol populations, respectively. These reanalyzed 9-month CT data will be cross tabulated to 36+ month CT data for presence of bone in Table 12.1 for the mITT population.

Finally, 36+ month digitized scout films for the CT (directly comparable to 36+ month plain x-rays) will be compared to both 24-month plain film and 36+ month CT scan results in Table S.12.2 for the mITT population.

If the null hypothesis of the primary endpoint is not rejected and thus non-inferiority with respect to this endpoint is not established, consideration would be given to performing non-inferiority tests for each component of the primary endpoint.

## 5.6 Safety Analysis

Safety for this study will be assessed principally based on the examination of: 1) 'ongoing' adverse events at the time of the last Pivotal IDE study visit; 2) new serious adverse events; 3) new medical conditions/physical finding since last IDE study visit; 4) occurrence of any retreatment to the original treated spinal level; and 5) neurological status.

### 5.6.1 Adverse Events

Treatment-emergent adverse events to be collected will consist of new medical conditions/physical findings and new serious adverse events. All new treatment-emergent adverse events will be coded using MedDRA coding dictionary. Events will be determined to be serious based upon the investigator's assessment, as recorded on the CRF. Additionally, there will be follow-up on adverse events, which were ongoing as of the time of the Pivotal study completion

Detailed listings of ongoing adverse events, treatment-emergent serious adverse events, and new medical histories/physical findings will be presented in the listings.

#### *Ongoing Adverse Events*

The number and percentage of the ongoing AEs that were previously reported as SAE for each treatment group will be tabulated in Table S.16.1. Current AE status (ongoing, resolved) along with action taken since last visit for every event will also be summarized for each treatment. The action taken since last visit will be categorized as 'none', 'spine surgery', 'non-spine surgery', 'medication', and 'other'.

Any ongoing AEs that have become serious since the Pivotal IDE study will also be reported in the treatment-emergent SAE tables.

The number and percentage of patients experiencing ongoing adverse events at the last study visit for the Pivotal IDE study, and the number and percentage of events will be summarized by SOC, and by preferred term for each treatment in Table S.16.2.

#### *Treatment-emergent Serious Adverse Events*

The number and percentage of patients experiencing new treatment-emergent serious adverse events, and the number and percentages of events, will be summarized by SOC and by preferred term for each treatment group in Table S.17.1. The number and percentage of patients experiencing such events will also be summarized by SOC, preferred term, and severity for each treatment group in Table S.17.2, and by relationship in Table S.17.3.

Table S.17.4 will tabulate treatment-emergent SAEs by SOC and preferred term by duration of time since PFL until retreatment (<12 Months, 12-24 months, 24-36 months, 36-48 months, 48-60 months, and 60+ months).

Similarly to the ongoing adverse events, the outcome and action taken for the new treatment-emergent serious adverse events will be summarized in Table 17.5.

All SAEs will be reported in the treatment-emergent SAE tables.

#### *Treatment-emergent Medical History or Physical Findings*

The number and percentage of patients experiencing new medical histories or physical findings, and the number and percentages of such events, will be summarized by SOC and by preferred term for each treatment group in Table S.18.1. The number and percentage of patients experiencing such events will also be summarized by SOC, preferred term, and severity for each treatment group in Table S.18.2, and by relationship in Table S.18.3. The outcome and action taken for the new medical histories or physical findings will be summarized in Table 18.4. Any new medical histories/physical findings that are classified as serious are also reported in relevant SAE tables.

#### **5.6.2 Analysis of Retreatments**

The following will be tabulated for retreatments in Table S.19.1:

- The number and percent of patients who have a retreatment subsequent to S01-01US
- The type of retreatment
- Hospitalization status, and whether the event was performed inpatient or outpatient

Additionally, Table S.19.2 will tabulate these data by duration of time since PFL until retreatment (operative period, 24 months, 36 months, 48 months, and 60+ months).

Details of all surgical events other than retreatments will be presented in data listings only.

#### **5.6.3 Neurological Status**

The neurological status of each patient will be summarized preoperatively (baseline), and at the 36+ month follow-up visits. Shifts in status (normal, abnormal) from baseline to the post-baseline time points will be examined in Tables S.20.1 to S.20.4 by treatment group for muscle strength, reflexes, straight leg raises, and sensory evaluation, respectively. Chi-squared or Fisher's exact test will be used to test the difference in status between the OP-1 Putty group and the autograft group, as appropriate. Additionally, shifts in status from baseline to post-baseline time points within treatment group will be tested using McNemar's test within each treatment group.

The normal/abnormal status for muscle strength, reflexes, straight leg raises, and sensory evaluation is defined based upon the following hierarchies ("abnormal" trumps "not evaluable" which trumps "normal"):

Muscle strength – separately for hip, knee, ankle, and toe:

- Abnormal: "Absent", "Trace", "Poor", or "Fair" is entered for any of the segments (flexion, adductors, extensors) for either side.
- Not Evaluable: At least one of the segments is missing for either side.
- Normal: "Good" or "Normal" is entered for all segments for both sides.

Reflexes:

- Abnormal: "Decreased" or "Absent" is entered for any of the reflex segments (right knee jerk, right ankle jerk, left knee jerk, and left ankle jerk).
- Not Evaluable: At least one of the four reflex segments is missing.
- Normal: "Normal" is entered for all four reflex segments.

Straight leg raises:

- Abnormal: "Positive (pain)" is entered for either or both legs.
- Not Evaluable: Status is missing for either leg.
- Normal: "Negative (no pain)" is entered for both legs.

Sensory evaluation:

- Abnormal: "Impaired" or "Absent" is entered for any of the sensory segments (L3, L4, L5, and S1).
- Not Evaluable: At least one of the four sensory segments is missing.
- Normal: "Normal" is entered for all four sensory segments.

## 5.7 Additional Analysis

Additional analysis will also be conducted on the following measurements:

- Visual Analog Scale Results for Pain Assessment
- Donor Site Pain (autograft patients only)
- General Health Survey (SF-36)

### 5.7.1 Visual Analog Scale for Pain Assessment

The visual analog scale for pain will be summarized preoperatively (baseline) and at the 36+ month follow-up visits.

Descriptive statistics will be presented for actual value and change from baseline to the post-baseline time points for each treatment in Table S.21. Difference in change from baseline will be examined using the two-sample t-test to test a difference in mean between treatment groups. Additionally, change from baseline will be examined using one-sample t-test to test the mean change within each treatment group.

### 5.7.2 Donor Site Pain

The donor site pain will be summarized at the 36+ month follow-up visits for autograft patients only. The donor site pain will be rated using both visual analog scale and pain status (none, mild, moderate, severe).

Descriptive statistics will be presented for the visual analog scale for each time point in Table S.22.1 for autograft patients only. Pain status will be summarized by frequencies and percentages in Table S.22.2 for each category.

### 5.7.3 General Health Survey (SF-36)

The General Health Survey Scale will be summarized preoperatively and at the 36+ month follow-up visits.

If a subscale of SF-36 is missing an item, then means of the items in the subscale for that patient will be used to impute the missing value. This will only be done if fewer than one-half of the items in the subscale are missing.

The SF-36 yields an 8-scale profile of functional health and well-being scores as well as psychometrically based summary measures – the Physical Component Score (PCS) and Mental Component Score (MCS). The 8 scales are:

- Physical Functioning Scale
- Role-Physical Scale
- Bodily Pain Scale
- Mental Health Scale
- Role-Emotional Scale
- Social Functioning Scale
- Vitality Scale
- General Health Perceptions Scale

Descriptive statistics will be presented for the actual value and change from baseline to the post-baseline time points for PCS and MCS, as well as for each of the 8 scales, by treatment, in Tables S.23.1 through S.23.10. The changes from baseline will be examined using Wilcoxon rank-sum test to test the differences in means between treatment groups.

## 5.8 Immunology

Serum samples will be analyzed for patients who were positive for anti-OP-1 antibodies at the 24-month follow-up visit, and for patients who did not complete the 24-month follow-up visit but were antibody positive at their last recorded visit.

Among this subset of patients, the number and percent of patients with neutralizing antibodies at 36+ months will be summarized in Table S.24.1 for each treatment group. Descriptive statistics will be presented for the titer result for each treatment at 36+ months in Table S.24.2. Patient profiles for patients with neutralizing antibodies will be presented in Table S.24.3. The following information will be summarized in each profile:

- Overall success, non-imputed at 24 months
- Overall radiographic success (overall success criterion #1), non-imputed at 36+ months

- Success based on Oswestry Disability Index (overall success criterion #2) at 24 months and 36+ months
- Success based on absence of retreatment (overall success criterion #3) at 24 months and 36+ months
- Success based on the absence of treatment-related serious adverse events (overall success criterion #4) at 24 months and 36+ months
- Overall neurological success (overall success criterion #5) at 24 months and 36+ months
- Neutralizing antibodies status at 24 months and 36+ months
- Antibody titer results over time plot at 6 weeks, 3-, 6-, 12-, 24-, and 36+ months
- Overview of new treatment-emergent serious adverse events [(system organ class, preferred term, days onset since operation, duration, serious (yes/no), and potentially immunologically-related (yes/no)].

Treatment-emergent potentially immunologically-related AEs are defined as the following:

Adverse drug reaction	Haematuria
Anaemia	Haemoglobin decreased
Angiopathy	Hyperkalaemia Infection
Arthralgia	Leukocytosis
Arthritis	Procedural site reaction
Arthropathy	Pyrexia
Blood potassium abnormal	Rach maculo-papular
Blood potassium decreased	Renal failure
Blood potassium increased	Systemic lupus erythematosus
Drug eruption	Thrombocythaemia
Drug hypersensitivity	Thrombocytopenia
Haematocrit decreased	

Success outcomes (overall success, overall radiographic success, success based on Oswestry Disability Index, success based on absence of retreatment, success based on absence of serious treatment-related adverse events, and overall neurological success) for patients with and without neutralizing antibodies during the current study will be summarized in Table S.23.4 by presenting the number and percentage of successes for each treatment group at 24 months and 36+ months, as applicable.

The number and percentage of patients experiencing new treatment-emergent serious adverse events will be summarized in Table S.11.5 by treatment group and neutralizing antibody status for the following time periods: any time point, operative-discharge date, 24-month, and 36+

month. The analysis will be repeated for new immunologically-related treatment-emergent serious adverse events in Table S.24.6.

## 6.0 REFERENCES

An HS, Simpson JM, Glover JM, Stephany J, "Comparison Between Allograft Plus Demineralized Bone Matrix Versus Autograft in Anterior Cervical Fusion," *Spine*, 1995; 20(20): 2211-2216.

An HS, Lynch K, Toth J, "Prospective Comparison of Autograft vs. Allograft for Adult Posterolateral Lumbar Spine Fusion: Differences Among Freeze-Dried, Frozen, and Mixed Grafts," *J of Spine Disord*, 1993;8(2):131-135.

Bernhardt M, Swartz DE, Clothiaux PL, Cronwell RR, White AA, "Posterolateral Lumbar and Lumbosacral Fusion with and without Pedicle Screw Internal Fixation," *Clin Orthop*, 1992;284: 109-115.

Boden SD, Monroe MA, Martin G, Hutton WC, "Posterolateral Lumbar Spinal Arthrodesis with rhBMP-2/HA-TCP Following Laminectomy in the non-human Primate," *Trans ORS*, 1998;19: 249.

Boden SD, Schimandle JH, Hutton WC, Chen MI, "The Use of an Osteoinductive Growth Factor for Lumbar Spinal Fusion," *Spine* 1995;20(24)2626-2632.

Brantigan JW, "Pseudarthrosis Rate After Allograft Posterior Lumbar Interbody Fusion with Pedicle Screw and Plate Fixation," *Spine* 1995;19(11)1271-1280.

Campbell's Operative Orthopaedics, Eighth Edition. AH Crenshaw, (ed.), Mosby , Boston, 1992.

Cook SD, Wolfe MW, Salkeld SL, Rueger DC, "Effect of Recombinant Human Osteogenic Protein-1 on Healing of Segmental Defects in Non-Human Primates", *JBJS* 1995;77A: 734-750.

Cook SD, Baffes GC, Wolfe MW, et al, "The effect of Recombinant Human Osteogenic Protein-1 on Healing of Large Segmental Bone Defects", *JBJS* 1994;76A: 827-838.

Cook SD, Baffes GC, Wolfe MW, et al, " Recombinant Human Bone Morphogenic Protein-7 induces Healing in Canine Long Bone Segmental Defect Model", *Clin. Orthop.* 1994;301:302-312.

Cook SD, Rueger DC, "Osteogenic Protein-1," *Clin Orthop*, Number 1996;324: 29-38.

Cook SD, Dalton JE, Tan EH, Whitecloud TS and Rueger DC, "In vivo evaluation of Recombinant Human Osteogenic Protein (rhOP-1) Implants as a Bone Graft Substitute for Spine Fusions", *Spine*, 1994;19(15):1655-1663.

Dall BE, Rowe DE, "Degenerative Spondylolisthesis : Its Surgical Management," *Spine*, 1985;10(7):668-673.

Fischgrund JS, Mackay M, Herkowitz HN, Brower R, Montgomery DM, Kurz LT, "Degenerative Lumbar Spondylolisthesis with Spinal Stenosis: A Prospective, Randomized Study Comparing Decompressive Laminectomy and Arthrodesis with and Without Spinal Instrumentation," *Spine*, 1997;22(24):2807-2812.

Herkowitz HN, Kurz LT, " Degenerative Lumbar Spondylolisthesis with Spinal Stenosis," *JBJS*, 1991;73-A(6): 802-808.

Jorgenson SS, Lowe TG, France J, Sabin J, "A Prospective Analysis of Autograft vs. Allograft in Posterolateral Lumbar Fusion in the Same Patient," *Spine*, 1994;19(18): 2048-2053.

Kant AP, Daum WJ, Dean SM, Uchida T, " Evaluation of Lumbar Spine Fusion," *Spine*, 1995; 20(21): 2313-2317.

Katz JN, Lipson SJ, Larson MG, McInnes JM, Fossel AH, Liang MH, "The Outcome of Decompressive Laminectomy for Degenerative Lumbar Stenosis," *JBJS*, 1991;73-A(6): 809-816.

Katz JN, Lipson SJ, Lew RA, Grobler LJ, Weinstein JN, Brick GW, Fossel AH, Liang MH, "Lumbar Laminectomy Alone or With Instrumented or Noninstrumented Arthrodesis in Degenerative Lumbar Spinal Stenosis," *Spine*, 1997;22(10): 1123-1131.

Kwiatkowski TC, Meyer RA, Gruber HE, Tabor JR, Murakami T, Howard BA, Wozney JN, Hanley EN, "Spinal Laminectomy with rhBMP-2 in the Canine: A Safety Study," *Trans ORS*, 1997;18: 189.

Magin MN, "Enhancement of Lumbar Vertebral Interbody Fusion by Human Recombinant Osteogenic Protein-1, OP-1, in a Sheep Model", American Association of Orthopedic Surgeons (AAOS), 1998, New Orleans.

Minamide A, Tamakit T, Kawakami M, Hashizume H, Sakatar R, "An Experimental Spinal Fusion with Wintered Bovine Bone and Recombinant Human Bone Morphogenetic Protein-2," *Trans ORS*, 1998;(19): 659.

Mundy DR, "Regulation of bone formation by bone morphogenetic proteins and other growth factors," *Clin Orthop*, 1997;323: 24-28.

Rompe JD, Eysel P, Hopf C, "Clinical Efficacy of Pedicle Instrumentation and Posterolateral Fusion in the Symptomatic Degenerative Lumbar Spine," *Eur Spine J*, 1995;4: 231-237.

Sandhu HS, "Spinal Applications for Recombinant Bone Morphogenetic Protein: Early Experimental Results," 1996.

Sampath TK, Rueger DC, "Structure, Function and Orthopedic Applications of Osteogenic Protein-1 (OP-1)." *Complications in Orthopedics* Winter, 101-107, 1994.

Spengler DM, "Fusion of the Lumbosacral Spine: An Excellent Treatment Option for Selected Patients with a Variety of Spinal Disorders," in Instrumented Fusion of the Degenerative Lumbar Spine: State of the Art, Questions, and Controversies, M Szpalski, R Gunzburg, DM Spengler, A Nachemson (eds.), Lippincott-Raven Publishers, Philadelphia, 1996.

The Spine, RH Rothman and FA Simeone (eds.), WB Saunders Co., Philadelphia, 1992.

Steinmann JC, Herkowitz HN, "Pseudarthrosis of the Spine," *Clin Orthop*, 1992; 284:80-90.

Yone K, Sakou T, Kawauchi Y, Yamaguchi M, Yanase M, "Indication of Fusion for Lumbar Spinal Stenosis in Elderly Patients and Its Significance," *Spine*, 1996;21(2):242 -248.

Zdeblick TA, "The Treatment of Degenerative Lumbar Disorders: a Critical Review of the Literature," *Spine*, 1995;20(24S):126S-137S.

Kim, Yongjung J. MD; Bridwell, Keith H. MD; Lenke, Lawrence G. MD; Rinella, Anthony S. MD; Edward, Charles II MD, "Pseudarthrosis in Primary Fusion for Adult Idiopathic Scoliosis: Incidence, Risk factors, and Outcome Analysis" *Spine*, 2005;30(4):468-474

Kim, Yongjung J. MD; Bridwell, Keith H. MD; Lenke, Lawrence G. MD; Rhim, Seungchul MD; Cheh, Gene MD, " Pseudarthrosis in Long Adult Spinal Deformity Instrumentation and Fusion to the Sacrum: Prevalence and Risk Factor Analysis of 144 Cases" *Spine*, 2006;31(20): 2329-2339