Trends-in-Medicine

April 2009
by D. Woods

Quick Pulse

Trends-in-Medicine
Stephen Snyder, Publisher
2731 N.E. Pinecrest Lakes Blvd.
Jensen Beach, FL 34957
772-334-7409 Fax 772-334-0856
www.trends-in-medicine.com
TrendsInMedicine@aol.com

FDA Panel Rejects Stryker's OP-1 Putty

Gaithersburg, MD
March 31, 2009

Under the cloud of a federal grand jury investigation of Stryker Biotech's marketing of OP-1, a spinal putty with a human device exemption (HDE) for certain lumbar spine fusions, the FDA's Orthopaedic and Rehabilitation Devices Advisory Committee rejected premarket approval (PMA) for OP-1 by a 7-1 vote. The panel had many concerns about the company's study designs; redefinitions of "overall success," including changing presence of bridging bone to presence of any bone as an endpoint; potential immunologic problems; post hoc analysis; and lack of data supporting efficacy and safety claims.

The company was unable to satisfactorily answer many panel members' questions on a wide range of subjects. At one point, a vocal panel member asked a company doctor to show him an area on a slide of a CT scan, then snatched the pointer away from the company doctor and gave him an anatomy lesson. The panel member said that what the company described as bone growth resulting from OP-1 was actually growing in the spinal canal.

Perhaps the biggest hurdle for the company was arguing its decision for changing an endpoint from "bridging bone" to "all bone." Panel members generally agreed that the only way to make the PMA approvable would be to conduct a new study. They said that the company would, at the least, have to design another fusion study comparing bridging bone to all bone, using CT scans. One panel member said that he could envision the company doing such a study in one year.

Background

OP-1 Putty is a three-component product containing device and drug components: bovine Type I collagen from demineralized bone, a recombinant human bone morphogenetic protein (rhBMP-7, called recombinant human osteogenic protein, or OP-1), and carboxymethylcellulose (CMC). The prepared material is placed alongside the vertebral bodies as a graft for posterolateral fusions. One box of the product contains a 2 oz. vial of sterile powder (1 g) consisting of 3.5 mg lyophilized OP-1 protein and ~1 g collagen, and a 10 ml vial containing 230 mg CMC. During surgery, the contents of the vials are combined and then mixed with 2.5 ml of sterile normal saline. Two boxes of the device are required for each patient – one for each graft site.

The device is used to initiate the cascade of events responsible for bone formation, including recruitment of mesenchymal stem cells and cell proliferation and differentiation into chondroblasts. The bovine collagen is used as a resorbable scaffold for new bone formation. The device only includes 75-425 µm because Stryker believes that particles outside of that range won't allow bone formulation.

CMC is used as a thickener to turn the material into OP-1 Putty. The putty is then used instead of bone graft; there are no other differences between the surgical approach and techniques typically used for uninstrumented posterolateral lumbar spinal fusion procedures.

OP-1 Putty has an HDE only for (1) posterolateral spinal fusion procedures in skeletally mature patients with lumbar spondylolisthesis who have already had a failed posterolateral spinal fusion or (2) as an alternative to autograft in recalcitrant long bone non-unions where use of autograft is not feasible and alternative treatments have failed for at least six months. HDE devices are used to treat conditions that occur in fewer than 4,000 patients every year. Stryker submitted its application to the FDA for OP-1 in June 2006 but pushed back the approval timeline.

A grand jury is looking into allegations (detailed in an 8-K filing on March 10, 2009) that Stryker illegally promoted OP-1 products and Calstrux (another bone repair product which was recalled in 2006), misbranded devices, and submitted false reports to the FDA regarding the number of patients treated with OP-1 under the company's HDE. The government has been investigating Stryker for years, and the grand jury investigation may kill any remaining hope the company has for OP-1 approval.

Two former Stryker sales representatives pleaded guilty in November 2008 and February 2009 to promoting OP-1 for indications not approved by the FDA, knowing that such off-label use could lead to dangerous complications. Last July, the FDA said that OP-1 products had been linked to potentially fatal complications when used in off-label cervical spine procedures. The complications included swelling of neck and throat tissue. Stryker also got several warning letters from the FDA about problems at its OP-1 plant.

Public Session

Dr. Jay Mabrey, the panel chair and a hip and knee replacement specialist from Baylor University Medical Center in Dallas, opened the hearing. Although the panel had scheduled two public sessions, there was only one public speaker.

Pamela Adams, representing the Orthopedic Surgical Manufacturers Association (OSMA), urged the panel to "focus deliberation on the product's safety and effectiveness based on the data provided." She lectured the panel on the definitions of reasonable assurance of safety and effectiveness and valid scientific evidence.

The Stryker Perspective

Stryker Briefing Documents

In its briefing documents, Stryker said that OP-1 "has equivalent clinical and radiographic outcomes" to a bone fusion procedure, called autograft, without the complications. The company said, "OP-1 Putty was demonstrated to be statistically non-inferior to autograft with regard to the modified overall success (47.2% for OP-1 Putty and 46.8% for autograft, p=0.025), demonstrating that OP-1 Putty is comparable to autograft in the important parameters of radiographic success, clinical success, and safety. Overall radiographic success at 36+ months was clinically comparable to autograft but was not shown to be statistically non-inferior…It is felt that 36+ month data represent a significant positive addition to the submission by providing longer-term safety and efficacy data than is typically available for spinal devices under consideration for approval."

Stryker defended its decision to combine data from the 36+ month radiographic assessments with the original 24 month clinical outcome assessments to develop a modified overall success assessment, "It is appropriate to substitute the radiographic success data obtained using CT scans at 36+ months for the original 24 month radiographic information because the original plain film assessment was flawed and did not accurately assess for the presence of medial bone formation that we now know is common with OP-1 Putty…It would be preferable to have all components of a composite measure assessed at the same timepoint. However, this was not possible …given that CT scans were not obtained at 24 months."

Stryker Presentation to the Panel

Stryker explained why it kept changing its study, insisting that OP-1 Putty is non-inferior to autograft, fulfills an unmet need, and should receive PMA approval. The panel had many questions for Stryker, including questions about immunologic data, gamma irradiation, and the rationale for the 36+ month study. On the whole, the initial questions were polite and not combative.

Dr. Julie Krop, vice president of clinical and regulatory affairs for Stryker, told the panel that OP-1 Putty is safe and effective and provides an unmet medical need for certain patients. She said that OP-1 is needed because degenerative spondylolisthesis is a common problem; iliac crest bone graft is the standard of care but has significant drawbacks, including increased pain and morbidity due to bone graft harvest; and there is suboptimal bone graft material in patients with osteoporosis, diabetes, and poor vascularity. She said, "Extensive preclinical studies have been conducted without any safety signals, and an identical product has HDE approval." She told the panel that 15,000 patients have been treated under the HDE and another 25,000 patients were treated with the drug product globally since 2001. She said that OP-1 is effective, saying that Stryker's pivotal trials showed "clinically comparable results" on six out of seven endpoints and a 26+ month follow-

up extension study showed it to be clinically comparable on all seven endpoints, using the more sensitive CT scan.

Stryker consultant Dr. Jeff Fischgrund – the OP-1 principal investigator, editor-in-chief of the ***Journal of the American Academy of Orthopaedic Surgeons***, and a spine surgeon at Beaumont Hospital – told the panel that alternative therapy is "truly needed" for some degenerative spondylolisthesis patients. Traditional therapy for these patients is decompression (no fusion), but Dr. Fischgrund said that patients treated with decompression alone have poor outcomes compared to patients treated with decompression + fusion. The goal of fusion surgery is to create a bony union across the involved vertebrae, resulting in stability and good long-term outcome. However, Dr. Fischgrund said that negative aspects of the gold standard procedure, iliac crest autograft, include time, pain, and complications resulting from bone graft harvesting. He said, "There is no approved BMP for primary lumbar fusions. Surgeons are actively looking for approved alternatives because the bottom line is that we want to decrease patient morbidity and improve their outcome."

Dean Falb, PhD, Stryker vice president of research and development, explained the chemistry of BMP and the basic biology of OP-1. He addressed concerns about sterilization procedures during the manufacturing process, calling the company's procedures "consistent and reliable." Dr. Falb described how the company came up with its dose selection, "a threshold concentration of OP-1 at a given volume is required to induce bone formation." He explained that the 1 mg/cc dose was the most effective dose in animal models and summarized:

- Clinical dose chosen above the threshold for bone formation.
  - Dose based on rat, rabbit, dog, and primate studies.
  - Clinical dose (1 mg/cc) consistently above threshold in multiple models.
  - Basis for selection consistent with other BMP filings.
- Spine fusion efficacy shown in multiple species and models.
- Autograft is radiopaque and apparent CT volume includes unincorporated graft.
- OP-1 is radiolucent, and all bone volume seen on CT is *de novo* bone.

Dr. Falb discussed safety studies which he said showed no developmental abnormalities were observed with OP-1. He then discussed immunogenicity assays, about which the FDA reviewers expressed concern in the pre-meeting documents. He said that the company's "reported immunogenicity results are accurate" including:

- Binding antibodies detected by ELISA.
- Assays were developed based on FDA recommendations and were validated to meet FDA guidelines.
- Neutralizing status evaluated by cell-based assays.
- Positives identified using a statistically-based cutpoint defined to allow for a 5% false positive rate.

He told the panel that OP-1 data show:

- **Efficacy** – seen in all models regardless of antibody response.
- **Safety** – no immune-related safety observations, and normal development occurs in the presence of antibodies.
- **Immunogenicity** – spontaneous BMP antibodies seen in 5%-10% of healthy individuals.

Dr. Falb summarized the preclinical studies:

- Dose selection based on multiple preclinical studies.
- Systemic toxicology studies show no adverse effects.
- Local toxicology studies show no significant observations.
- Safety pharmacology studies show no effects.
- Developmental toxicology studies show no significant abnormalities.
- Developmental immunization studies show no significant abnormalities.

Dr. Krop defended the pivotal trial data, arguing that OP-1 showed non-inferiority to autograft:

- All clinically relevant outcomes in the study were comparable between OP-1 Putty and autograft.
- Radiographic success rates using well accepted criteria were comparable between OP-1 and autograft.
- Safety outcomes were reassuring and comparable between OP-1 and autograft.
- OP-1 Putty patients had high rates of neurologic success at least as good, if not better than autograft. This last point was important, she said, because "we know other BMPs have been associated with neural complications."

The primary endpoints at 24 months were:

- Clinical success.
  - ODI – 20% improvement.
  - Neurologic – no clinically relevant neurologic changes.
  - Treatment-related serious adverse events – no serious adverse events related to treatment.
  - Re-treatment – no re-treatments intended to induce fusion.
- Radiological success.
  - Angulation – ≤5º of angular movement (using plain films).
  - Presence of bone – a marker of OP-1 activity.
  - Translation – ≤3mm of translational movement (using plain films).

Dr. Krop said, "We did not achieve our 24-month endpoint of overall success because the presence of bone was not detected…Why were all endpoints comparable except presence of bone?" Her answer: CT scans originally read for inter-transverse bone formation only were carefully re-evaluated.

- Scans were collected at nine months only and were not part of the original primary endpoint.
- Bone formation was present but was more medial than anticipated.
- Medial location of the bone formation led to under-estimation of presence of bone at 24 months. Readers instructed to look only for inter-transverse process fusion on x-ray.
- Plain x-rays provide poor visual ion of medial bone; vertebrae themselves may block the visualization of medial bone.

Dr. Krop told the panel that OP-1 Putty:

- Is safe and avoids the morbidity associated with autograft bone harvest. The safety profile was reinforced by data from 36+ month extension study and extensive post-marketing data.
- Achieved clinical success on all key clinical parameters. The success persists through 36+ months, which is a clinically more rigorous point.

She concluded, "Therefore based on the data…we conclude that this PMA is approvable."

Dr. Krop described the extension study which she said was designed "to correct for the insensitivity of plain films and assess the efficacy of OP-1 Putty compared to autograft using a more sensitive imaging modality: 36+ month CT scans with mean follow-up = 4.4 years." The extension study also collected 36+ month assessments of angulation, translation, and all clinical endpoints.

Dr. Krop said that the company contacted patients who had moved and enrolled 80.2% of them in the study, "(The primary endpoint at 36+ months was) the same as the original endpoint, except radiographic assessment used CT scans…in order to address the insensitivity of plain films at evaluating medial bone formation. The timepoint was also used to assess the re-treatment rate – the most significant clinical endpoint."

She said that the primary non-inferiority endpoint was achieved, with 47.2% overall success for OP-1 compared to 46.8% for autograft (p=0.025). She said that OP-1 and autograft were similar in all components of the "overall success." Using 36+ month data only, there was no statistically signficiant difference in radiographic and clinical outcomes between the two groups. Dr. Krop added that the two groups showed no statistical difference in several key additional endpoints: visual analog scale for right lower extremity at 24 and 36+ months and SF-36 physical function scores.

Dr. Krop summarized the safety data:

- OP-1 Putty is marketed under an HDE.
- More than 15,000 patients have been treated in the U.S. under HDE approval since 2004, and no trends in serious adverse events have been seen. An average of 0.28 adverse events reported per 100 units of OP-1 sold in the U.S.
- Additional 25,000 patients treated with OP-1 products worldwide – OP-1 Implant in U.S., Osigraft in Europe, Canada, and Australia.
- Pivotal trial safety profile consistent with postmarketing data.
- Safety of OP-1 Putty treatment in PLF is similar to autograft treatment with respect to the proportion of patients experiencing treatment-emergent and treatment-related adverse events, serious adverse events, neurologic complications, neoplasms, and deaths.
- 16 patients died during study participation – 5.3% in OP-1 group and 5.8% in autograft group.

Dr. Krop said that the incidence of antibodies peaked at three months and then declined significantly, "but the mean antibody titers returned to baseline by 24 months. In terms of neutralizing antibodies, the peak formation occurs at 3 months and by 12 months only 1 patient had neutralizing antibodies and beyond no patients had neutralizing antibodies…We carefully assessed efficacy…safety. In terms of all adverse events, there was no observed impact of neutralizing antibodies on efficacy."

She summarized the company's immunogenicity data, saying antibodies to OP-1 do not pose a safety risk to patients:

- 5%-10% of patients have antibodies against OP-1 prior to exposure.
- More than 40,000 patients globally treated with either OP-1 Putty or OP-1 Implant, and no safety signals related to immunogenicity have emerged.
- Pivotal trial patients evaluated for adverse events related to immunogenicity had no difference in adverse event profile in the two groups. Serum creatinine showed no differences from baseline.

Dr. Lee Katz, head of musculoskeletal radiology at Yale University Medical School, spoke for the company about radiological issues in the pivotal trial. He posed several questions and offered answers.

Why were plain films insensitive in the evaluation of the original endpoint?

- OP-1 produced bone formation more medial than anticipated. Readers were forced to visualize traditional lateral inter-transfer process fusion.

- Plain film technique may interfere with visualization of medial bone formation.
  - Plain films detect an average of densities from front to back in two dimensions.
  - Medial structures difficult to visualize as vertebrae themselves may block visualization.

Why CT Scan Results Were Inconclusive

Time period for presence of bone	OP-1 Putty	Autograft
9 months	80%	100%
36+ months	75%	77%

Why were the 9-month CT scan results inconclusive? Dr. Katz said 9 months is not an adequate timepoint to compare bone formation between OP-1 and autograft.

- Autograft is detectable immediately after surgery.
- Residual autograft is likely to be present at 9 months, and thus there is bias in favor of autograft.
- This is confirmed by the fact that there is only 77% of autograft patients with presence of bone by CT at 36+ months compared to 100% at 9 months.

Why gather additional data? Dr. Katz said that was done to correct for the inability to measure the presence of bone, which was the primary endpoint. He added that CT scanning is the gold standard in evaluating bony vertebral body anatomy as well as new bone formation, "The study endpoint used CT scans rather than plain films to allow for more accurate detection of new bone formation, especially medial." He concluded:

- CT scans allow for a more appropriate comparison between treatment groups.
- From a radiologist's perspective, the key endpoints for determination of successful fusion are stability by angulation and translation, coupled with good clinical outcomes.
- 36+ month radiographic assessment is comparable to what the results would have been at 24 months had CT scans been obtained.
  - 9 and 26+ month CT scan results comparable for OP-1 Putty.
  - Autograft rates appear to decrease but likely due to artificially high rate at nine months.

Dr. Huub Schellekens, professor of pharmaceutical biotechnology at Utrecht University in the Netherlands, described the antibody response in OP-1 Putty. He told the panel that there is "no observed impact of OP-1 immunogenicity on the clinical efficacy and safety of the OP-1 Putty product." He said that most therapeutic proteins are immunogenic and that there "is no correlation between a high incidence and clinical consequences." He said that Stryker's reported neutralizing antibody data were accurate and that the company worked with the FDA to develop a new ROS (reactive oxygen species) neutralizing assay which was used in the pivotal study. Dr. Schellekens said that immunogenicity is multifactorial, and potential causes for OP-1 immunogenicity include product irradiation, protein aggregation, protein dose, and route of administration. He said that non-irradiated OP-1 Putty is immunogenic in primates, "Human and primate OP-1 amino acid sequences are identical and non-irradiated OP-1 induces antibody production in primates."

Dr. Schellekens summarized the cause of OP-1 immunogenicity:

- Immunogenicity is multifactorial.
- Non-irradiated OP-1 was immunogenic in preclinical studies.
- Route of administration and dose appear to be relevant.
- Mechanism is breaking tolerance.
  - Direct interaction with B-cells.
  - No memory – minimal concern for re-treatment.

He gave an overview of the development concerns:

- OP-1 is a single use product.
- Neutralizing antibodies are transient in patients, decline rapidly after three months, and are gone by 24 months.
- There is no preclinical evidence of immunogenicity effects on development.
- Labeling requires female patients to avoid pregnancy for one year; this will be incorporated into physician training.

Dr. Schellekens concluded, "There is no effect of immunogenicity on efficacy in preclinical models, and efficacy is independent of neutralizing antibody status in pivotal trial patients…My overall risk assessment is a reasonable assurance of safety. I'm also in a position like yours in the Netherlands, and I ask, 'Is the product safe? Would I use it?' And my answer here is yes."

Eugene Poggio, PhD, president and chief biostatistician of Biostatistical Consulting, described the various study designs and summarized:

- In the original study, analysis of primary endpoint did not demonstrate non-inferiority.
- It is evident, however, that detection of bone by plain film in the study was biased in favor of lateral bone and hence in favor of autograft.
- Extension study using CT scans to detect bone was conducted to rectify the issue.
- Results combining clinical results from original study and radiographic results from extension study were thought to be less biased and more meaningful due to use of CT scans.

- The results were robust and consistent.
- Two treatments have virtually identical estimated overall success rates.
- Based on overall success, OP-1 is, to a statistical certainty at most 11.6% worse than autograft, and at best 12% better than autograft.
- Two treatments are similar regardless of method of handling missing data, analysis population, and even variations on primary endpoint.
- Two treatments are also very similar across each of seven components of overall success.
- The totality of evidence supports non-inferiority of OP-1 compared to autograft.

Dr. David Wong – study investigator, an orthopedic surgeon, director of the Advanced Center for Spinal Microsurgery at Presbyterian St. Luke's Medical Center in Denver, and past president of the North American Spine Society (NASS) – talked about understanding data in a clinical setting. He asked, "What are the important outcomes for clinicians and patients? Does it work? Does it improve quality of life? Is it safe?" He answered that OP-1 Putty fulfills an unmet clinical need because there is no other approved product for posterolateral lumbar fusion, and OP-1 Putty avoids iliac crest bone graft harvesting. He said that OP-1 has an excellent safety profile, and it works, "Overall, it's effective, and it's safe."

THE FDA PERSPECTIVE

FDA BRIEFING DOCUMENTS

In briefing documents given to the panel before the meeting, the FDA reviewers expressed doubts about the efficacy of OP-1 and questioned its safety, specifically in the area of immune response. The FDA reviewers criticized Stryker's analyses of data, questioned the company's changing definitions of overall success, and expressed concern over lack of data requested by the FDA but never received.

In the area of efficacy, the FDA reviewers said that three different analyses of OP-1 showed that it failed to perform as well as the control. The reviewers also questioned a fourth analysis that the company said showed non-inferiority of OP-1 to control. The FDA reviewers criticized how Stryker formulated the trials and kept changing the definitions of success, and the reviewers said that the data submitted by the company are in question. In addition, the reviewers were concerned about the possible long-term immunologic effects from antibodies developed in response to OP-1. The reviewers also said that they have concerns about the manufacturing of OP-1, though the panel is not going to discuss those concerns.

The reviewers wrote, "The agency has questions whether or not this combination product can be considered relatively safe and effective when compared to the autograft control. Based upon the currently submitted pivotal study data, the source of the effectiveness differences between the subjects treated with OP-1 Putty vs. the control treatment is not known and is left as a point of discussion for the panel. It might be due to selection of the current human dose for the OP-1 component; due to the differences in the ability of the product to promote bone formation in humans in contrast to the behaviors observed in lower order animals; due to changes in the potency or stability of the recombinant protein after exposure to irradiation sterilization; or due to a clinical response to the oxidized, truncated, and aggregated protein."

The reviewers said that there are safety questions associated with the immune response to OP-1, "It is not clear if the observed antibody response (antibody rates, types, and duration of effect) is due to the terminal radiation processing or some inherent characteristic of the protein. The Agency has questions concerning this as well."

Non-clinical studies

The FDA reviewers had "numerous concerns" about OP-1 because it is a combination product containing biologically-derived and device components, for example, immunological response or ability to signal tissue formulation, dosage, etc.

Non-clinical pharmacology/toxicology studies

BMP-7 (the active agent in OP-1) has an important role in the development of the skeleton, nervous system, eye, kidney, heart, and germ cells. It also has a role in tissue repair or regeneration in the kidney and brain.

- ***Pharmacology:*** Animal studies showed that OP-1 Putty was as effective or superior to autograft in enabling spinal fusion. No human dosing studies were done, so the optimal human dose is unknown. The FDA reviewers asked "whether the observed level of effectiveness of the product resulted from the selection of an improper dose or some other factor."
- ***Toxicology:*** Toxicity studies in mice and rats did not raise significant concern. Studies on the effect of OP-1 on tumor cell proliferation were inconclusive.
- ***Immunogenicity:*** In the pivotal clinical trial, the OP-1 immunogenicity was demonstrated by a high incidence (94%) of anti-OP-1 antibodies, including antibodies with OP-1 neutralizing activity. In many patients, the antibodies were still significant up to 24 months postop. The FDA reviewers said that neutralizing antibodies can interfere with endogenous BMP-7, and "this presence over time is of particular concern for women of childbearing potential."
- ***Reproductive toxicity:*** The reviewers said that animal studies showed a possible risk for prenatal developmental effects in women treated with OP-1 products.

Immunogenicity

Stryker initially developed four assays to evaluate OP-1 immunogenicity. One of them, a ROS cell line-based neutralizing assay, was "determined to be irreparably flawed," according to the FDA reviewers. Stryker developed another neutralizing assay, but the collected sera were not re-analyzed, and the actual rate of neutralizing antibodies is still unknown and may "be different from the reported rate of 25.6%."

Chemistry, manufacturing, and control concerns

The FDA reviewers said that the main concern with the chemistry, manufacturing, and control data is over the terminal radiation sterilization of the product. Stryker terminally sterilizes OP-1 using high levels of gamma irradiation (24.5 to 31.5 kGy), and the effects of radiation on protein structure and function, especially at such a high level, are unknown. Although protein oxidation, aggregation, and truncation occur, there are no data showing whether the reported antibody response rate is related to it.

Although Stryker was asked to analyze data from *in vitro* biological activity and do additional assays, no data were provided to the FDA. Stryker gave the FDA a brief data summary and several graphs from some non-clinical studies and a current IND study, but it did not submit complete test reports, and FDA reviewers "could not fully correlate the results from those non-clinical evaluations to observed results from clinical trials."

The FDA reviewers had other concerns that will not be part of the panel discussion, including:

- The need for an appropriate calculation of release and stability specifications for the OP-1 drug substance.
- OP-1 extraction efficiency from the OP-1 Implant.
- Adequacy of the removal of adventitious organisms from the collagen matrix.
- Adequacy of control over parameters affecting the percentage truncation of the OP-1 molecule during fermentation.
- The determination of protein yields for each process step for the OP-1 drug substance.
- Lack of an assay at release and stability measuring acid content.

Clinical studies

Stryker did three clinical studies: a pilot study, a pivotal study, and an extension study. None showed non-inferiority of OP-1 to control. The pivotal study was the main source of clinical data, and the extension study was done to address specific FDA questions. The pilot study's design was so different from the pivotal study that it was not used to evaluate safety and efficacy.

The FDA reviewers said that the pivotal study did not show that OP-1 was non-inferior to autograft in terms of overall success rate in any of the success definitions, including the one which excluded radiographic data and which would have been potentially biased in favor of OP-1. The reviewers wrote, "The overall success rate in the control group was higher when compared to the investigational group and was primarily due to the higher rate of overall radiographic success in the control group. The lack of bone formation, especially bridging bone, appears to be the primary source of the poor radiographic outcome in the investigational group. Control subjects had a higher rate of bone formation compared to the OP-1 subjects at three months postop, and this remained higher throughout the remainder of the follow-up period."

Before Stryker submitted its data for the PMA to the FDA (after all clinical data had been collected but before the database was closed), the company submitted a modification that changed the definition of overall success and the definition of the efficacy populations for analysis. It also modified the fixed non-inferiority margin from 10% to a maximum of 14%.

As part of its PMA submission, Stryker included a post hoc analysis. It offered a second change of the overall success definition and removed all the radiographic data from the predefined composite primary endpoint. Stryker said that it made the change because "evidence for presence of bone on x-ray was inconsistent with the beneficial effects noted in angular and translational movement." Success was now defined as:

- Oswestry Disability Index (ODI) success.
- Absence of re-treatment.
- Absence of serious treatment-related adverse events.
- Neurological success.

This analysis did not use data from all the patients nor did it use a sensitivity analysis to look at the potential impact of the exclusion. The reviewers had these concerns:

- OP-1 patients are expected to form a solid fusion mass in the absence of any bone graft. Bone formation relies solely on the ability of the recombinant protein signal to elicit cellular differentiation and proliferation. Short of another invasive surgery to look for a fusion mass, the FDA was "unaware of another reliable way to detect the presence of bone other than by radiographic techniques."
- Stryker's rationale for the post hoc analysis – inconsistency between the bone formation and spinal stability – "was made without supporting data." Although the FDA asked for an analysis assessing a correlation between bone formation and spinal stability, it did not receive one.
- Stryker's post hoc analysis was biased in favor of OP-1 because:
  - The primary endpoint was redefined retrospectively.

- The radiographic component was the sole blinded endpoint.
- A much higher percentage of the excluded subjects were from control.
- Most of the excluded control subjects were considered successes at 12 months postop.

➢ The Type I error rate could be seriously inflated without any adjustment for the post hoc nature of changing the primary endpoint and its associated multiple analyses. The reviewers said, "All of these issues were expressed to the sponsor. Even acknowledging the questions, the agency reviewed and analyzed the data…and determined that, using the original definition of overall success approved in the clinical study, both the intent-to-treat analysis and the per-protocol analysis failed to show that OP-1 was non-inferior to autograft and in fact showed that the control treatment was superior to OP-1."

Overall Success Rates with OP-1

Time	Control n=87	OP-1 n=208	p-value for non-inferiority
Original protocol (pre-specified statistical analysis plan)			
3 months	42%	20%	0.971
6 months	37%	30%	0.340
12 months	48%	31%	0.863
24 months	48%	32%	0.824
Per protocol analysis			
3 months	47%	22%	0.982
6 months	43%	32%	0.553
12 months	60%	31%	0.990
24 months	57%	35%	0.942

Efficacy and Safety of OP-1

Endpoint	Months	Control n=87	OP-1 n=208	p-value for non-inferiority
ODI	12	84%	83%	Nss, 0.45
	24	85%	80%	Nss, 0.180
No re-treatment	12	96%	95%	<0.001
	24	93%	9%	0.003
Safety				
No serious treatment-related adverse events	12	96%	91%	Nss, 0.078
	24	96%	89%	Nss, 0.141
Neurological adverse event	12	97%	98%	<0.001
	24	94%	100%	<0.001
Radiographic fusion	Control was significantly superior to OP-1			
Radiographic results				
By original success definition	3	53%	25%	Nss, 0.997
	6	48%	37%	Nss, 0.543
	12	61%	40%	Nss, 0.961
	24	65%	40%	Nss, 0.989
By per protocol analysis	12	73%	38%	Nss, 1.000
	24	74%	40%	Nss, 1.000
By second radiographic success definition	12	73%	49%	Nss, 0.985
	24	75%	52%	Nss, 0.961
	24 (MI)	69%	53%	Nss, 0.622
By second per protocol analysis	12	74%	49%	Nss, 0.990
	24	75%	52%	Nss, 0.960
Postop follow-up period				
Presence of bone	3	83%	52%	Nss, 1.000
	6	87%	60%	Nss, 1.000
	12	73%	49%	0.985
	24	75%	52%	0.961

The FDA reviewers said that even using the second definition of overall success, where the radiographic endpoint was modified, OP-1 still failed to show non-inferiority to control. They added that the overall radiographic success rate was lower for OP-1 vs. control, "This large difference in overall radiographic success rate was primarily driven by the differences in bone formation (bridging bone in accordance with the original radiographic success definition compared to any bone formation in the second radiographic success definition)." In the control arm, 83.1% of patients had presence of bone at 24 months vs. 61.7% of patients with OP-1 ($p<0.001$). Using the third definition of overall success, which excluded all of the radiographic data, the FDA reviewers said that Stryker "claimed that (OP-1) was statistically non-inferior to the control treatment…(but) later agreed with the Agency that such claim…was not appropriate."

In the extension study, a post hoc re-analysis of the 24-month plain films and 9-month CT scans on a subset of treated subjects, Stryker said that the original radiographic assessment was erroneous because the 24-month plain films in many cases did not show evidence of bone formation, while the nine-month CT scans did. The reviewers wrote, "The sponsor stated that two points could be derived from this evaluation. The first was that the use of plain films was inappropriate for the evaluation of their product. The second was that the initial radiographic reviewers had been looking in the wrong location for the fusion mass." The new extension study was designed to collected a single CT scan from any available subjects that could have been taken anywhere from 36+ months to almost 72 months postop. Using the data, Stryker proposed a fourth definition of overall success, which combined the 24 month clinical outcome data with the new 36+ month CT scan/re-operation data. Stryker said that the data showed that OP-1 treatment was statistically non-inferior to the control treatment.

The FDA reviewers had these concerns about the extension study:

➢ The post hoc nature of the re-analysis that determined that original data were faulty.

➢ Evidence of implant migration, which had not been previously observed.

➢ Different follow-up timepoints for different elements of the success definition.

- Reliance on longer-term data "that is not consistent with our understanding of clinical practice."
- The post hoc re-definition of the primary endpoint.
  - The confirmatory nature of the original pivotal trial was compromised.
  - A statistical non-inferiority claim could not be made.

Safety

The FDA looked at two types of analyses: a traditional analysis that would be applicable to any product used in spinal fusion, and an analysis looking at potential immune responses to the recombinant protein and bovine collagen components of OP-1.

Traditional safety analysis

The FDA reviewers said there was no significant difference in adverse events between OP-1 and control. However, they said the OP-1 arm had a higher numerical incidence of serious adverse events. They wrote, "This is consistent with the finding that the investigational (OP-1) treatment was ***inferior*** to the control treatment with respect to the success rate based on the absence of serious treatment-related adverse events."

Immunological safety

Because OP-1 contains a recombinant human protein which could elicit an immune response, Stryker was required to collect serum from all subjects and perform assays looking for the presence of OP-1 antibodies and neutralizing antibodies. In the pivotal trial, 94% of patients tested positive for anti-OP-1 binding antibodies. In addition, although antibody titers decreased over time, many patients still had significant titers even out to 24 months postop. The reviewers said that they were concerned about the long-term effects of the antibodies, "The agency believes that the high incidence and long duration of anti-OP-1 antibodies raises questions because OP-1 has been demonstrated to have important roles in fetal development as well as adult mammals."

FDA reviewers said that the assay was unreliable and that the panel will have to discuss the issue. They said that they had asked Stryker for more information about immunological response to OP-1 but said, "Based on their response or lack of responses to these requests, it is not clear that the sponsor has provided adequate information to address the immunological safety questions the agency has regarding the product."

FDA Presentation to the Panel

The FDA reviewers criticized Stryker's pivotal and extension clinical studies, including definitions of overall treatment success, the study designs, statistical concerns, and safety and clinical efficacy results. The reviewers said that, regardless of the definition of treatment success, OP-1 Putty was not found to be non-inferior to autograft in the treatment of single level degenerative spondylolisthesis (Grade 1-2) in patients undergoing decompression and uninstrumented posterolateral lumbar fusion.

The reviewers said that OP-1 Putty was not shown to be non-inferior to autograft in overall treatment success as prospectively defined at the beginning of the pivotal study, (definition #1) and after subsequent revision of the definition of success (definition #2). Although immunogenicity did not appear to play a role in adverse events in OP-1 patients, there was a trend towards decreased overall treatment success and radiographic success in patients who developed neutralizing antibodies compared to those who developed non-neutralizing antibodies.

The FDA reviewed previous FDA action concerning Stryker's OP-1, including a letter outlining multiple deficiencies, such as:

- Key safety issues not adequately addressed.
- Did not meet primary endpoint (overall subject success at 24 months) approved in original IDE.
- Did not meet revised endpoint proposed in pre-PMA submissions.
- New issues resulting from additional revised primary endpoint provided in response to major deficiency letter.
- Inadequate responses to concerns associated with manufacturing, potency, dosing, and immune response.

Chemistry, manufacturing, and control (CMC) concerns

The FDA said that a major concern with the CMC data is over the terminal radiation sterilization of the product. Ionizing radiation is an effective method for eliminating microorganisms including bacterial and viruses. It is used for surgical instruments and devices as well as some pharmaceuticals and foods. 25 kGy is the recommended dose to sterilize medical devices, and the OP-1 Implant is sterilized with 24.5 kGy to 31.5 kGy. However, gamma irradiation is not typically used for biologic (protein) drugs, due to their general sensitivity to the effects of ionizing radiation. The typical sterilization method used for biologics are filtration and aseptic processing. The direct effects of ionizing radiation on proteins include breakage of covalent bonds randomly along the polypeptide chain, causing protein truncation and inactivation. Larger molecules are more susceptible. Indirect effects include oxidation, damidation, disulfide modification/shuffling, and cross-linking. Observed changes induced by gamma irradiation on OP-1 protein and Implant include loss of activity (30% decrease in potency assay after extraction from OP-1 Implant), aggregation (around 19-fold higher increased levels of OP-1 aggregates), increased amounts of truncated and oxidized variants, increased immune response to OP-1 Implant, and development of neutralizing antibodies against OP-1 Implant and potential cross-reactivity on endogenous BMP-7.

The reviewers summarized:

- γ irradiation is used to sterilize OP-1 Implant.
- γ irradiation is not used for approved recombinant protein products.
- γ irradiation causes loss of biological activity, aggregation, truncation, and oxidation of recombinant human OP-1.
- A high incidence of immunogenicity is observed with γ irradiation OP-1 Implant.

Immunogenicity

Concerns about anti-OP-1 antibodies on endogenous BMP-7 include the fact that no data were provided to the FDA regarding antibody cross-reactivity.

Concerns for Antibodies in the Clinic with OP-1

Clinical concern	Clinical Outcome
Safety	Neutralize activity of endogenous counterpart with unique function causing deficiency syndrome Hypersensitivity reactions
Efficacy	Enhancing or decreasing efficacy by extending or decreasing half-life Decrease efficacy by altering biodistribution away from target
Pharmacokinetics	Antibody production may dictate changes in dosing level due to PK changes
None	Despite generation of antibodies, no discernable impact

The reviewer summarized:

- There was a high incidence of binding (94%) and neutralizing (25.6%) antibodies developing in patients treated with OP-1 Putty.
- 41% of subjects still tested positive for binding antibodies 24 months post-treatment.
- The impact of these antibodies on the long-term health of those patients is not understood.
- No patients tested positive for neutralizing antibodies after 12 months.
- 36.7% of subjects tested positive for binding but not neutralizing antibodies at 36 months.

The FDA reviewer said that the issue is "not a show stopper," but it relates to patient health. As for aggregates and immunogenicity, she said that aggregated proteins tend to be more immunogenic than their non-aggregated counterparts and that protein aggregation may qualitatively and/or quantitatively impact the immune response. She concluded that there is insufficient data regarding OP-1 to understand the impact of aggregates on immunogenicity.

The reviewers said that they requested modified protein manufacturing to address concerns associated with gamma irradiation, potency, and stability. They also asked for new non-clinical and clinical dosing studies and a clinical trial. Additionally, the FDA asked for more manufacturing information, improved antibody assays, and an additional reproductive/toxicology study.

Dr. Ryan Kretzer from the FDA's Division of General, Restorative, and Neurological Devices, Office of Device Evaluation, Center for Devices and Radiological Health (CDRH), described the clinical studies: the pilot study, the pivotal study, and the extension study. The pilot study showed that OP-1 looked promising, but the autograft treatment group showed the highest percentage of patients with bridging bone formation. OP-1 Putty showed high pseudoarthrosis and immunogenicity rates compared to control.

The revised definitions of overall treatment success in the pivotal trial were acknowledged **but not approved** by the FDA. CT imaging was performed on all patients at nine month post-treatment in order to assess for bridging bone formation and pseudoarthrosis, but this was NOT included as a criteria for patient success or as a study endpoint. Dr. Kretzer said that:

- OP-1 was not shown to be non-inferior to autograft in:
  - Overall treatment success (using success definition #1 or #2).
  - ODI success.
  - Radiographic success (using either success definition).
- OP-1 was shown to be non-inferior to autograft in:
  - Absence of treatment.
  - Neurological success.

Safety results showed similar rates of adverse events and deaths in the two groups. Although not statistically significant, there was a trend towards a higher rate of treatment-related serious adverse events in the OP-1 arm compared to control (12% vs. 7%, p=0.22).

OP-1 Safety Results

Measurement	OP-1 Putty only	OP-1 Putty + Autograft
Pseudoarthrosis	42% *	25%
Immunogenicity: antibody titers at 6 months	92%	83%
Neutralizing antibodies at 6 weeks **	29%	0
Pseudoarthrosis in patients with neutralizing antibodies	57%	N/A

* 30% of patients required re-operations
** 57% of patients who develop neutralizing antibodies also experience pseudoarthrosis

Overall treatment success, using definition #1 (approved by the FDA) at 24 months, was defined as a composite of:

- ≥ 20% improvement in the Oswestry Disability Index.
- Radiographic spinal fusion.
  - Bridging bone on x-ray at the treated level *and*
  - ≤ 5 degree angulation on flexion-extension x-rays *and*
  - ≤ 2 mm translational motion on flexion-extension x-rays.
- Absence of a decrease in neurological status (muscle strength, reflexes, sensory, straight leg raise) unless attributable to a concurrent medical condition or to the surgical procedure.
- Absence of re-treatment.
- Absence of treatment-related serious adverse events.

The FDA presented 9-month CT scan data, which Stryker did not include in its PMA.

9-Month CT Scan Data

Measurement	OP-1 Putty	Autograft
Any bone formation	85%	99%
Bridging bone formation	1%	54%

The FDA said this was important because any bone means just that – bone volume that may not provide any structural support, compared to bridging bone formation, which does provide support. Stryker had changed its definition of overall success by changing bridging bone to any bone formation. At 36+ months, 56% of OP-1 patients had bridging bone formation compared to 83% of autograft patients (p=0.001).

Jianxiong (George) Chu, PhD, an FDA biostatistician, critiqued Stryker's study hypotheses and design, saying that, "Concerns over the sponsor's claim of non-inferiority based on their post hoc analysis and the analysis of the extended study show a Type I error rate inflation and were probably biased in favor of the OP-1 Putty group…According to the pre-defined 10% non-inferiority margin, the sponsor's modified intent-to-treat (mITT) analysis, (with or without imputation for missing data) of the extended study still failed to support the non-inferiority claim even without any adjustment for the retrospective change of the primary endpoint."

Dr. Chu said that Stryker proposed a revised statistical analysis plan (SAP) in December 2005 when the study was nearly completed. The overall radiographic success was changed to:

- Presence of bone (rather than bridging bone).
- Angulation of ≤ 5 degrees.
- Translational movement of ≤ 3 mm (rather than ≤ 2 mm).
- The fixed non-inferiority margin of 10% was modified to be variable, ranging up to approximately 14% depending on the success rate in the control group.
- The efficacy population for analysis was changed into an mITT analysis which included all treated patients with at least one post-treatment follow-up visit.
- For the overall success and overall radiographic success endpoints at 24 months, missing data imputation was changed from LOCE to multiple imputation.

Dr. Chu said that the FDA reviewers had concerns with late-stage changes:

- Significant changes (primary endpoint, non-inferiority margin) were proposed by the sponsor when the study was close to the end. Be aware that this is an open-label study.
- The sponsor's proposal to allow a larger non-inferiority margin is not justified from a statistical point of view since the close-to-maximum variability was already accounted for in the original sample size estimation, which assumed near 50% overall success rate for both groups.

He argued that Stryker's own post hoc analysis was flawed. As for the extension study, he said that there was too much missing data. Dr. Chu said, "The sponsor's imputation accounts for 30% of total treated patients, and underlying statistical assumption for – in this case – such an assumption may not hold because the patient could be doing well, and not going back to participate in the study."

In the extension study, Stryker said that the CT 36+ months data showed a success rate of 74.8% for the OP-1 arm and 77.4% for the autograft arm. However, Dr. Chu said that missing data or non-evaluable data were excluded. A re-evaluation of the data by the same statistician that evaluated the data for Stryker showed 80% success for the OP-1 arm and 100% success for the autograft arm, "From my view, most sensitivity analysis will go down to just this issue…Based on my analysis, the sponsor not only excluded the missing data for the analysis…(but) ignored all the missing data due to the other reasons."

According to the original protocol-defined SAP and the revised SAP, OP-1 Putty was not shown to be non-inferior. Dr. Chu said, "We have concerns over the sponsor's claim of non-inferiority based on their post hoc analysis, which they concede, and the analysis of the extended study. The two concerns are Type I error rate inflation and also the probably biased (data) in favor of OP-1 Putty group."

Panel Questions for FDA Reviewers and Company Experts

Dr. Paul McCormick, a neurosurgeon at Columbia University, said that he had some "real concerns" about medialization, "It was not identified as an issue at all in the pilot studies. In fact, numerous pictures showing what I assume to be a robust… process of fusion – were shown. So I have some concerns regarding if medialization was an issue, why didn't we see it earlier in the pilot study?"

Immunologic Response

David MacLaughlin, PhD, a biochemist at Harvard Medical School, asked about sterilization, "Why didn't you use sterile filtration and avoid the irradiation at all? And I want to discuss the endogenicity story as it relates to the possibility of either re-treatment, which I think is prohibitive, but I'm still thinking about the consequences of that, and whether people screened for hypersensitivity would be eliminated." A speaker for Stryker said that the company thought that it was essential to bind OP-1 to the surface of the collagen during manufacturing, "There are other BMPs on the market, and those are much higher doses. For safety and efficacy we thought it essential to combine the two."

Dr. Raj Rao, a spine surgeon at the Medical College of Wisconsin, praised Stryker for its presentation and congratulated the company "for a very honest study." But he had some questions generally focused on the immunologic issue, "There seems to be a discrepancy in the information you gave us…Antibodies were found in 25% of patients at 36 months, and some of the other speakers have said it returned to baseline at 24 months. The baseline based on that study appears to be approximately 7.95% for OP-1 and can you clarify that discrepancy."

A Stryker expert said that they were measuring two different things, "One is patients with any presence of antibodies and the other is the mean titer rate – measuring the titer itself – so it's more of a cutoff point issue…There still are 25% of patients above the statistical cutpoint." Dr. Rao asked if changing the labeling would be sufficient. He then asked about the pivotal study, "You mentioned that you used well accepted criteria in the pivotal study. It seems to me that well accepted criteria would be bridging bone. The bibliography you provided doesn't talk about the presence of bone, it talks about bridging bone, and I wonder why your scientific published materials looked into the bridging bone and reported 56% for OP-1 and 83% in the autograft at 36+ months. Why didn't you include that data in your PMA and why did you choose not to use the presence of bridging bone? It seems to be the identification of medially located bone may be more suitable for a study that's looking at the efficacy of OP-1 on the osteopathic process than in creating fusion?"

Dr. John Kirkpatrick, a spine surgeon at the University of Florida College of Medicine, asked about Stryker's statement about memory in the immunology studies, "I didn't see the data that would state that. Second, once an antibody is produced by the body, the cells that produce that can be ramped up to cloning a lot faster than they were before. A couple of questions might help me. Of the patients who had the antibody at the start of the study, was their success rate different from those without antibodies at the start of the study? You reported 5%-10% of presence of antibodies pre-control in the study. Do you have data to show that the second use of the OP-1 device has any different outcomes from a first use? Either in animal or human? It was mentioned that it is second use, but there may be a result in a second issue of spondylolisthesis…Does it make any difference from the first?"

Dr. Janine Jason, an immunologist from Hilton Head Island SC, had several questions about Stryker's slide presentation and asked for answers about specific slides. She also asked if there were any studies on T-cell reactivity and OP-1 and collagen.

The consumer representative asked about the immunogenicity and BMP-7 and whether kidney function was tracked throughout the study. An FDA reviewer said that the HDE does not require patients to be followed on a set schedule.

Dr. Jason asked about the neutralizing assay, and the FDA reviewer responded, "The FDA was not satisfied with the original neutralizing antibody assay...The sponsor later developed an assay that we were satisfied with…The early samples were not re-tested but in the new assay the later data which were negative were tested using the new assay…The early data are not data that I consider reliable. The later data showing no effect at much later timepoints are probably much more reliable data."

Dr. Jason asked the company what data it has about how much of the OP-1 leaves the site over time. She also asked if all 40,000 people who have received the product had received irradiated product. A Stryker spokesman said that the products were the same, and an FDA speaker objected, "No that's not true. There was an amendment to the IDE last year…a major amendment to the IDE application last year, and we determined that they were highly similar, but they are not the same, and the process is no longer similar to the PMA process. But they are gamma irradiated. The manufacturing has changed." Stryker representatives nodded their heads no and said that they would argue that point later.

The panel chair, Dr. Mabrey, had two questions, "What I have been hearing is concern over the irradiation because it tends to create these protein aggregates which lead to formation of antibodies. Do the antibodies bind to the active dimmer – as they do to the protein aggregate, and what is the clinical effect? Second, given this patient population, could the sponsor outline what percentage of the population was female plus what percentage was of childbearing age, and how many of those became pregnant, and any follow-up on that?"

Dr. MacLaughlin asked the FDA reviewers about the impact of radiation on the extracted protein and its biological activity, "Was that done in comparison with non-radiated comparison material?" The reviewers answered yes. Dr. MacLaughlin said, "It's clear that these proteins – when purified – can find aggregates, truncation, so it's coming with the recombinant material. That is important to recognize in the product…The blocking antibody question is…difficult to assess…You have to block 100% of the protein present…I'm suggesting the presence of antibodies might be more significant than the presence of neutralizing antibodies. It's also important to ask if there are any data about the PK effects of the neutralizing antibody population vs. people without." An FDA reviewer said, "The non-irradiated protein is greater than 97% pure, so the issue of truncation and aggregation and oxidation are much reduced in a non-irradiated protein compared to irradiated. So your assertion is true, but probably at a much reduced rate." Dr. MacLaughlin answered, "It's a matter of degree. Where does the risk show up?"

Second dose response

A Stryker expert said of the memory effect or second dose effect, "In bone formation models, we have not done repeat dosing, and that is not our intended application…it's a contra-indication."

Later in the discussion, Dr. Kirkpatrick had a sharp exchange with a Stryker official that caused spectators to gasp:

- *Dr. Kirkpatrick:* "Is a second use a contraindication? Did you say that?"
- *Stryker official:* "It's a warning."
- *Dr. Kirkpatrick:* "We don't need to bog down on this. It'll be a condition if it goes that far…Say I do a fusion and five years later the patient comes back with a similar situation. Does that patient have any different outcome risk/pseudoarthrosis risk than the first dose?"
- *Stryker physician*: "The immune response would be the same. It's not like a vaccine reaction."
- *Dr. Kirkpatrick:* "You're telling me that there is no consequence of having an antibody at the beginning or having an antibody as a result of treatment?"
- *Stryker physician:* "There were some patients treated who were not positive at pre-treatment and at follow-up. Re-fusion would be far enough out that the antibody would not be affected."

Commenting on the discussion, Dr. MacLaughlin said that the antibody would have to be recognized the second time. He asked if any patients known to have antibody going into the study changed in titer and fusion. A Stryker official said that there were only eight patients in the study. She added that Stryker was not recommending that it be used a second time.

Asked about the biology, an FDA reviewer said, "We're not just worrying about patients being treated a second time, we don't even know whether a rise in BMP would be enough to cause a rise in immune response…There's a lot about the biology of this protein that we don't know yet, and we need to be investigating clearly and carefully. That is not necessarily something to make one say, 'No, don't approve this product.' The amount of information we don't have needs to be understood when thinking about this product in terms of efficacy."

Statistics

Dr. Kirkpatrick asked, "The ODI measure was 20% improvement. You had patients varying from 30 to 100 on the scale; that means we could have a six point difference for those who started at 30, or a 20 point difference for those who started at 100, in theory. The literature reports a difference of 12 points, not percent. Why did you choose 20%, and why would that result in minimal clinical difference in your patients?" A Stryker speaker said, "If you look at the literature there is a wide range – from 15 to 20 points to 20% – the literature has evolved over time. We remained consistent, which was the percentage. We can get the numbers for you."

Panel member Brent Blumenstein, PhD, an independent biostatistician, had many questions for Stryker. He said, "I assert that I don't agree with the way it was done because the trial computations appear to be based on a control arm success rate of 50%, but the variable margin was based on the extreme ends of the variable margin scale…I'm questioning your rationale for the way you did the 36-month extension. You claim that the reason you did it was because you felt there was an under-ascertainment of bone formation…at 24 months, based on a methodology that missed the 24-month formation of bone. I guess you brought patients in, used CT for control and for intervention arm patients, and compared for non-inferiority between the control arm and investigational arm patients based on the 36-month CT measurements. It seems to me that what you should have done is to use the 36-month CT measurements in the investigational arm as a correction for the measurement done at 24 months and then go back and compare the data at 24 months. This interests me, because you say that the formation of bone in the control arm patients would degenerate in time. You're setting yourself up for a comparison where the control arm will have less success with respect to bone formation as measured at 36-month CT scan. I'd like to see the comparison of the 36-month bone in the investigational arm vs. 24 months in the control arm. I think that you might be doing something that's not quite right by using the 36-month measure in the control arm." A Stryker official said that it would take several weeks to do that comparison, and the panel chair quashed that idea.

Kathleen Propert, ScD, a statistician at the University of Pennsylvania, asked about people who didn't make it to treatment, "I know about the 295 people who made it to treatment, but I'd like to know about the 41 who were randomized but didn't get that far…Why is that? There are a lot of places where the

dropout between the two arms was different. This is another case of understanding how the imputation was done and the assumptions used for that would affect dropouts. Also, there was a hint, not statistically significant, that there was a difference in treatment-related serious adverse events between the two groups…I'd like to know (the data) in both groups."

Multiple imputations

A Stryker expert – a public health consultant from Harvard – tried to answer the question about multiple imputations, saying that there was no difference in the analyses and outcomes between analyses that made use of multiple imputations and those that did no imputations or used other imputation approaches. He said, "The characteristics of the people who returned for the 36+ month follow-up were similar to the entire eligible population. Multiple imputation is a well accepted, commonly used procedure. It produces valid results …under broader observations…Clinical success or radiologic success were modeled separately. Important predictors were included in making the prediction of the missing outcomes, but in every case only a single covariant was found to be important. The main point is that our analyses were similar if we did stratified analysis, radiographic success, multiple imputation, or no imputation at all."

Non-inferiority margin change

A Stryker official said, "We felt that a 10% change from a 10%-20% success rate was very different from a 40%-50% success rate. We took it to our statisticians to attempt to address these concerns." A Harvard biostatistician, speaking for Stryker, asked the panel not to worry about a 10% margin.

Why is there no impact on Type I variability?

A Stryker physician said, "If you look at the study, in my opinion, the primary endpoint is still the same. It's just a matter of the measurement – two were changed from a less precise, less sensitive to more precise and more sensitive (measures)…Type I error probability won't be sacrificed. We should not be penalized by using a more precise measurement."

Any T-cell reactivity studies with collagen and BNP complex?

A Stryker expert said that there is no effect of OP-1 on T-cell activity.

Pharmacokinetic activity

A Stryker official said that more than 95% of the protein is gone from the site at 35 days. The blood half-life in primates is less than one hour. He said, "It initiates a cascade early in the process and then is rapidly removed from the blood." Dr. MacLaughlin said that "We're looking at migration of radiographic material and it leaves the implant site? So then the extraction issue is one of biopotency not of availability? You get it all out, it's just less active."

HDE and PMA relationship

The industry representative asked about the standard for determining safety for the HDE compared to a PMA. An FDA spokesman said, "You look at safety for that specific patient population – a risk/probably benefit ratio." He also asked about the percentage of antibodies in humans. Asked if labeling could be changed to reflect that, the FDA immunogenicity expert said that the FDA "needs more and better data in order to complete the labeling."

Dr. Kirkpatrick asked about process, "We have a product and all I can see is a difference in use. We are now reconsidering it as a PMA. We have already discussed the safety portion of it. But a lot of questions are coming up about safety. Does a PMA have to stand on its own, or does a previous approval as an HDE product change anything about the process of approval?" An FDA official answered, "In general, a product has to stand on its own. But an HDE is approval of a product for a different indication of use. Questions here are based on the data in the PMA and not necessarily from the HDE…In an HDE we don't necessarily see any clinical safety data. It's a discussion of how does the product work, what is the proposed mechanism of action, can it be based on animal data, theoretical information, and how does that match up with what you could expect to see in a safety profile…In the HDEs we had some clinical information from a different population but not a complete safety analysis…Here is a product that could potentially help them because we can make a probable benefit argument." Dr. Kirkpatrick asked, "So, practically and fundamentally, they are different?" The FDA official said, "Yes, different definitions of safety."

A Stryker physician answered a question about the HDE, "The HDE has to be renewed annually. In order for me to renew it, I need to fill out a form and state whether there have been any serious adverse events. Assuming everything goes well, they will renew it. I'd imagine that it would not be renewed if there were any serious adverse events." He then discussed the dropout rate. "I'll take part of the hit for this. I helped design the study. Back then we told the patients to which arm they were randomized. When you discuss studies with patients they …get disappointed when they don't get into the investigational group. I hate to say that the dropout rate in the control arm was higher because patients were disappointed…If you look at the two groups, these patients were randomized but did not undergo treatment, there is no difference between the two groups. They're in their 70s, and the male to female ratio is similar, but if you look at the amount of angulation, translation, and ODI scores, they're all very similar. There is no difference in the demographics between the two groups." Dr. McCormick asked, "Would you agree that that creates the placebo 'nocebo' effect? Patients believed that they would have less pain and they were pleased with being assigned to the treatment, and that added to the intensity of care brought by surgeons committed to the product. The 'nocebo' patients were disappointed, and they were not going to do as well. The surgeon wouldn't have as high an intensity in terms of the treatment of the patient."

Asked how many dropouts had surgery outside, the Stryker physician couldn't answer the question. Dr. Kirkpatrick said, "We don't know why they dropped out. They might have gotten better. I don't think the number of dropouts was a big deal. I think it was just random chance that it happened."

Another FDA reviewer reiterated the difference between PMA and HDE safety information, "There is a different amount of information that needs to be approved in a PMA compared to an HDE…The product was HDE approved for the use described, for the very narrow patient population, and you can't extrapolate what may be safe from the HDE population to the more general population. Given the conditions of an HDE and how it differs from a PMA, we made a decision that the OP-1 product was approved for the specific orphan population, fewer than 4,000 patients per year in the U.S. Those products should be relatively safe and relatively beneficial – not that they ***are*** relatively safe and beneficial. We don't have data to submit (re the HDE) because it's whatever the company happens to hear about. We may be missing information, and we don't know what those datapoints might be. We never receive the information because it is not required to be collected." The FDA panel member said, "When you're looking at a risk:benefit ratio of safety and effectiveness, you weigh in the population. It is approved as safe and effective for that limited indication for which there were not alternatives…This product is trying to be an alternative to autograft, which is a reasonable alternative, but they're trying to address the benefit of a second surgical site."

Bone

Dr. Propert, a statistician, said, "I need to understand better the difference between bridging bone and total bone in terms of statistically significant." Dr. McCormick told Dr. Propert, "Imagine a bridge made of bricks. They can be put together so that they can cross a river and support your car, or they can lie in the river and make a dam. You have bone growing the two vertebrae together, or you can have a bunch of bone just sitting there and which is ineffective in immobilizing the two segments."

A neurosurgeon speaking for Stryker tried to answer questions about the differences between medial and lateral bone formation and how that impacts clinical outcomes. A reviewing editor for the article in the ***Journal of Neurosurgery, Spine,*** said that the lateral flexion and extension radiography "is really the only reliable determinant to assess whether fusion has occurred." He talked about the angulation success and translation success, which were relatively equal between autograft and OP-1 in the 24-month data (instead of presence of bone, which favored autograft), "Patients will not do well in long-term follow-up if you don't have a successful biomechanical procedure. As a clinician, I see a product that seems to have good long-term outcomes based on disability index scores, it appears to be safe, associated with good neurological outcomes." Dr. McCormick asked about the trend toward neurological success, "I don't see any reason for it." The Stryker physician said, "There are some important take-home points. Bone morphogenic proteins have the potential of causing diverticulitis. We're not seeing adverse neurological impact. The second point is whether, in fact, talking about whether medial or lateral bone is better or worse, the possibility may be there. Could, in fact, the long-term fusion results be more robust? The third possibility is that it's just random chance."

A Stryker physician said, "The amount of bone that is formed is clearly sufficient to stabilize the spine. I'm convinced that if OP-1 did not do its job the patients would not be doing well at four and a half years. There is criticism that we had to extend the study, but the fact that I have long-term data makes this a better study for me to show how the patients do. Yes, the bone is not where we thought it would be, but it's stabilizing (the patient). I have 4.5-year data, and the patient is doing well…and at the end of the day that is what I want to see."

Radiologic success definition

A Stryker consultant answered the question of why medial bone was not described in the initial pilot study. He said that bridging posterolateral bone was seen on plain films on 78%, which was sufficient to proceed with the pivotal study. He added that only 12 patients were in the study. As preclinical studies were evaluated, "We looked back and did notice occurrence of medial bone formation. Once the medial bone was identified, CT scans were employed for the clinical study. We also characterized the stabilizing effect of medial bone relative to lateral bone and found comparable stabilizing effect." Dr. Kirkpatrick interrupted the physician and asked him to prove what he had just said. The Stryker speaker said that biomechanical testing (of animals) showed no significant differences in medial vs. lateral bone in both arms. He said, "Both preclinical and clinical data showed no significant difference in the stability animals or patients with medial or lateral bone formation." Dr. Kirkpatrick asked for the evidence, but the Stryker speaker did not have it. Dr. Rao asked if the Stryker speaker was talking about bridging medial and bridging lateral bone. The speaker said, "These were animals deemed to be fused based on the biomechanics; most were bridging bones." Dr. Kirkpatrick explained to the rest of the panel what they were talking about, "The medial bone is the bone growing closer to the…axis. My concern is that they have not verified that for us in the data."

Adverse events

A Stryker official answered a question about treatment-related adverse events, saying that the FDA had reported that the rate of treatment-related serious adverse events was not statistically significant. She said, "The percentage of patients with serious adverse events was similar: 50% OP-1 Putty to 49.4% autograft. Twelve percent of patients had treatment-related serious adverse events in the OP-1 Putty group and 6.9% in the autograft group (p=0.1868). There are some things you

expect to see, like pseudoarthrosis. There were some cardiac (events), minor things that don't seem to be related, and for autograft they look very related to product."

An anatomy lesson

Dr. Kirkpatrick asked a radiologist or spine specialist to use a laser pointer to outline the medial facet border on a slide with the patient on an axial view.

- *Stryker physician:* "This is a reformatted image based on the axial view…It'd probably be in this area here."
- *Dr. Kirkpatrick:* "I would advocate a different interpretation. May I borrow your pointer? *(takes pointer)* The natural facet point is here. The bone is here, lateral, medial. And if one followed medially, one sees either inadequate compression or bone into the canal – one of the two. Please comment."
- *Stryker doctor:* "It's hard to assume."
- *Dr. Kirkpatrick:* "I acknowledge that. That's what we've heard all day. Thank you. Some of your surgical colleagues may differ from that. We're not talking about bone formation medial for the facet point. That's desirable. We're talking about…against the lateral aspect of the facets. My concern is that it appears we have bone formation in the canal. I don't know if it's bone formation from the OP-1…But medial is a relative term, and it does not mean medial to the facet joint. *(another collective gasp from the audience)*."

Dr. Mabrey, the panel chair, asked if the FDA had any clarifications to any questions. An FDA official said "The real issue is that if this drug is considered to be effective then we will work with the company to mitigate risk and understand the risk and deal with it. It is a risk. It is an unknown risk. We have questions about the efficacy of this product, and that has to be taken into any risk:benefit assessment. Should the product be considered to be effective, we then work with the sponsors to mitigate those risks."

FDA Questions to the Panel

Question 1. Irradiation sterilization. The combination product is provided sterile after exposure to relatively high levels of gamma irradiation (i.e., 24.5-31.5 kGy). Based on Stryker's data, this induces numerous changes in the recombinant protein, including oxidation, aggregation, and truncation. These changes to the protein likely contribute to the observed high incidence of anti-OP-1 antibodies in subjects receiving the product (94% of investigational subjects), including the development of antibodies that neutralize OP-1 activity (26% of subjects).

Comment on the potential for changes in the recombinant protein, including oxidation, aggregation, and truncation to have an impact on:

- **The stability or potency of the recombinant protein component of the combination product.**
- **The biological activity of OP-1 Putty.**
- **The immunological response to the combination product, and clinical effects that ensue from such responses.**

The panel chair summarized: "The panel generally believes that the stability of the product is maintained after irradiation. It generally believes that the bioactivity is retained in the presence of the irradiation, and it generally shows some concern over the immunogenicity of the product as a whole. Suggestions have been made for possible screening of potential patients."

Comments by panel members included:

- *Dr. McCormick, a neurosurgeon:* "While there are some concerns in the long term, I didn't see any adverse safety issues based on the data presented by the sponsor."
- *Dr. MacLaughlin, a biochemist:* "I believe in general that when you purify a recombinant protein, it is not the same as the endogenous material. Then the problem becomes, what is your standard of evidence for efficacy. If you irradiate it, changes are going to happen, especially at these high doses. We don't have enough data to assess how much. Damage is done, but retains biological activity. Also, true recombinant proteins are antigenic in humans. One would argue that the radiation could increase the antigenic capacity of the protein, although we have one study which said it wasn't true. But *in vivo* there could be significant changes induced in the half-life of the protein…so I think that's an issue. So I'm willing to concede there is damage done, but material retains potency. When it comes to the issue of immunological response and subsequent events, I am concerned. I'm not comforted by or persuaded by the data which we look at the binding antibody vs. neutralizing antibody. I tried to do that…but it's exceedingly difficult to show…so not showing a correlation to response doesn't convince me. Other measures need to be made for that. I am concerned.

I think that the antigenicity is a question, but we have data presented here that it remains effective. We just don't know if this is the optimal response or not. The radiation is an issue. I know that it was selected for reasons of efficacy in making the product, but I'm concerned, going forward, that there is an issue and people might need to be screened before they get the product."

- *Dr. Kirkpatrick, a spine surgeon:* "The biologic activity doesn't appear to make a significant difference clinically …I continue to have questions in my head about the protein changes. I also have a question about 'least burdensome.'"
- *Dr. Jason, an immunologist:* "The information presented on potency and biological activity are convincing. The data on immune safety are reassuring, but I'm still concerned about the possibility of rare events. We know that that kind of processing with cellulose could open up endogenetic sites that aren't normally seen. It's conceivable that there could be rate complications, and so I am not completely reassured."
- *Dr. Rao, a spine surgeon:* "I have continued concerns on the immune response to the protein considering OP-1 in fetal development. I'm not sure that we have the answer to that."
- *Mark Melkerson, director of the FDA's Division of General, Restorative, and Neurological Devices, Office of Device Evaluation, CDRH:* "Regarding least burdensome, this is a combination product, and I've heard it called a combination product device. It is a combination product, but in terms of the different components…the regulations for a drug component are still for a drug component."

QUESTION 2. Definitions of overall success and statistical analyses. During the PMA review, Stryker proposed four definitions of overall success and made three major modifications to the statistical analysis plan (SAP) before database lock. The original pivotal study showed that OP-1 is significantly inferior to control in terms of the primary endpoint. The non-inferiority claim was still unsupportable under the late-stage revised SAP. Stryker conducted the extension study with a new primary endpoint and concluded that non-inferiority was demonstrated.

Comment on:

- **The clinical soundness of the various definitions of overall success**
- **The statistical soundness of Stryker's claim of non-inferiority**

The panel chair summarized: "There are serious concerns about the clinical soundness of the various definitions of overall success. There are problems with the multiple imputation mode, post hoc analysis, as well as the introduction of Type 1 errors."

Comments by panel members included:

- *Dr. Propert:* "I have one comment, which is that I am a little concerned about patient expectations. This was an unblinded study, and I'm worried about some of the other biases going on – they maybe showing up in that. One blinded endpoint is the one that is potentially the most controversial, and it doesn't make me more comforted that the patient reported outcomes are dependent on more objective endpoints. I have major concerns about the biases in this study, and – just for the benefit of the panel – multiple imputation is appropriate, but under certain assumptions that I really am not sure hold here, having to do with why data are missing, and accounting for that in the process. There have been dropouts from the beginning. Finally, as to the population available for the extension study…I don't see any way a statistical analysis can bring more light on the outcomes."
- *Dr. MacLaughlin:* "I wondered why CT wasn't part of the protocol in the first place."
- *Dr. Kirkpatrick:* "At the time the study was designed they were sound endpoints…CT technology improved at that time and shortly after that, so it wasn't in the purview. Adding it at the end brings up issues of statistical soundness. As one trying to understand the issue…it's drummed into us that post hoc analysis is not supposed to be taken into account…so from the standpoint of answering this question, I have significant concerns about the statistical soundness. I think the clinical soundness was fine at the beginning of the study."
- *Dr. Jason:* "What leaves me concerned is the FDA reanalysis which suggested that people in the control group were not represented. I have significant concerns about that."
- *Dr. Rao:* "From a statistical standpoint, a non-statistician standpoint, I have some concerns about the dropouts, the change in the endpoints over time, unblinding the study before the final endpoints were determined. From a clinical standpoint, the process of determination of the final radiographic endpoint and the CT scan – if the spines were stable in flexion extension, there must be some fusion mass somewhere. But we know, based on prior studies, that you can have fibrous unions without bone in the absence of a fusion mass, so I'm not sure that it's an entirely valid extrapolation that there must be some bone somewhere. Also, if CTs were selected as an endpoint, which I think is reasonable at the 36-month point, instead of choosing the presence of medial bone, we should have chosen the presence of bridging bone somehow…These are the concerns I have with the development process of the endpoints."
- *Dr. Blumenstein, a biostatistician:* "What we have here is an abuse of alpha or Type 1 probability or whatever you want to call it. They changed the margin without taking into account that they kept the trial size the same. They did a post hoc analysis and made a decision to proceed

with the 36-month analysis, for example. The consequence is that the p-values that you see aren't interpretable as p-values in the typical FDA regulated setting. You need to take these p-values as being just measures of strength of evidence, and you have to keep in mind that what I said about alpha having been abused and much larger than declared...The sponsor is asking you to take the collection of data they presented...and to use your clinical judgment as to whether this combination is efficacious."

- *Dr. McCormick:* "I still have some concerns…The sponsor published in peer reviewed journals…and in December 2008 there was also a pivotal trial. He quoted the journal saying that the company proceeded with a pivotal study, but did not reproduce the results of the pilot study and the primary endpoints. He said, "Medialization of the bone graft, to me, would have shown the problem. I saw nothing to suggest a medialization problem in the animal studies. There was a change in the subcomponent, not to bridging bone on CT but to any bone. The problems I have are numerous. First, it was an ad hoc analysis. The presence of bone has never been suggested as an indication of fusion in any study I've ever seen. How do we know that bone is new, and how do we quantify it?… The real inference of solid fusion here is related to the fact that there was comparability in the two groups with respect to the outcome of angulation and translational motion…The study population was weighted heavily. Also, it is well known that patients with stenosis – with minimal flexion extension – routinely do well with surgical decompression alone. I don't think it's fair to use angular translational data in this particular population because they were so stiff…The idea that we would analyze these in a post hoc fashion is also problematic."

QUESTION 3. Clinical performance – effectiveness. Comment on the clinical effectiveness of the combination product. Include the potential necessity for performing a human dosing study to assess the correlation between the reported effectiveness and selection of the correct dose of the recombinant protein component of the combination product.

The panel chair summarized: "The panel generally believes that, at best, the product's clinical effectiveness is equal to that of autograft, but there are several concerns with regards to the nature of the patient population at the beginning of the study and the statistical analysis of whether the studies showed or demonstrated non-inferiority. With regard to dosing, it's felt that at the beginning dosing was reasonable, but the sponsor may want to look at dosing density. One of the main concerns is about the nature of the patient population as it was selected and in the unblended nature of the study itself."

Comments by panel members included:

- *Dr. MacLaughlin:* "From the dosing point of view, the rationale was pretty good at the beginning. But having seen it be completely antigenic in the subjects, adjusting the doses upward doesn't make a lot of sense."
- *Dr. Kirkpatrick:* "On the dosing issue, we don't know how they got the results they did…There may be a regional dose response or dosing density per square centimeter or whatever you want to do, looking at where that product lies in order to get good fusion. There are other questions about why they got the varied results that they did. As for a new clinical study, that's a huge question. You might be able to sort out the dosing in an animal model. It would be extremely challenging to do the same model. The question is, would we find clinical differences other than radiographic ones…in the long run? My impression is that with the radiographic concerns, the application of whether that makes a solid fusion – I'm not clear that it creates a fusion."
- *Dr. Rao:* "Clinical effectiveness is largely independent of the product – the product is aimed at a fusion after decompression. As far as pain relief, there doesn't seem to be any significant difference between this and autograft. In terms of radiographic effectiveness, the PMA data with the presence of bone alone is difficult to interpret. However, if we were to use the published peer review literature, it shows inferior effectiveness in the OP-1 group. I'm not sure that a new dosing study is feasible, and I will defer on a new clinical study."
- *Dr. Blumenstein:* "Now that I learned that patients started off in almost success, or a non-differentiating state, I'm concerned that what may have been found might have been non-inferiority to something that wasn't working very well."
- *Dr. McCormick:* "There's no question that in many patients the product worked extraordinarily well. The problem is predicting who will have that response and who won't…I think the challenge is finding out who will benefit the most from this. Whether a dosing study is a variable associated with patient response, I don't think that is required of the sponsor."
- *Dr. Propert, a statistician:* "I'm not convinced that there is evidence of non-inferiority."

QUESTION 4. Clinical performance – safety. Comment on the safety of the product. Include the potential for clinical concerns associated with the immune response to the recombinant protein including any that potentially could affect either maternal and child health.

The panel chair summarized: "It is generally believed that the device is safe, however, the panel has expressed significant concerns with respect to the immunogenicity, especially with the fact that it crosses the placenta. There have been some concerns about the size of the study population and whether it is large enough."

Comments by panel members included:

- *Dr. Kirkpatrick:* "They didn't show that there was a safety concern…with regard to the immune response, those have been brought up several times…I still have my reservations. As far as the maternal/fetal barrier, we've heard that the antibodies can cross…and that is a valid concern. With regard to general immune issues, we have to be concerned that we've only looked at 300 patients. We have a drug standard to be dealing with and that may require a much higher standard of safety."
- *Dr. Jason:* "The data are reassuring as far as safety...but there are still questions."
- *Dr. Kirkpatrick:* "I had to live through patients on Vioxx (Merck, rofecoxib)…I don't think that HDE or the world-wide database has enough to show that it's not there."
- *Dr. McCormick:* "I think the data show that the product is safe. There will be long-term concerns about material fetal long-term reaction and delayed response that will only be answered with more data. The burden is on the sponsor to show that it is safe."

Before the vote, Dr. Wong summarized Stryker's position, "First consider that what we're applying for is a situation of clinical unmet need. As well, in terms of the situation with instability, back in 1999 the base of techniques was not widespread. Even though patients were in a relatively stable state, decompression was clearly a destabilizing situation. Finally, in terms of the selection bias, one of the strengths we as clinicians see is the long-term 4.4 year outcome data, which stays similar throughout the whole course. The treatment or placebo effect is seen early to potentially a year or two. But the treatment effects have stayed the same out to 4.4 years in the extension study."

Dr. Jason asked if conditions can include new studies. The FDA panel member said that would be in the realm of the non approvable in order to make a decision on safety or efficacy. Dr. McCormick asked if limiting the indications for use would be an appropriate condition. The FDA panel member said that a subanalysis of existing data could be done.

QUESTION 5. What is the panel's recommendation on approval of the PMA for OP-1 Putty?

VOTE: 7 to 1 not approvable

Dr. Blumenstein moved that the PMA not be approved. Dr. Propert seconded the motion. Dr. McCormick was the only dissenting vote. Panel comments on their votes included:

- *Dr. Blumenstein:* "I am unconvinced that the data provide sufficient evidence of efficacy because of the flaws in the study design and the abuse of probability."
- *Dr. MacLaughlin*: "It seems safe. I don't feel commenting beyond that, but the material as it's used seems to be safe to me."
- *Dr. Rao:* "It sounds reasonable."
- *Dr. Kirkpatrick:* "I want to make sure that the members of the panel understand the vote. From a clinical standpoint, having an ability to not do the graft harvest is an important aspect of our considerations, but whether that is outweighed by the concerns we have regarding efficacy and safety."
- *Dr. Jason:* "In terms of immunologic safety, I don't think we have enough data. I think some simple studies could be done that could be reassuring, but right now the data are not adequate. You could not do a subanalysis. You could use the patient population…but it would have to be a new study."

After the vote, the panel members stated the reasons for their votes.

- *Dr. Rao:* "My concerns are the lack of radiographic efficacy and the choice of the presence of bone instead of bridging bone and the lack of clear non-inferiority of OP-1. I have no major concerns with regards to the safety issue. I do have some concerns about maternal/fetal transportation of antibodies and concerns about fetal development."
- *Dr. Jason:* "I have concerns…about potential study bias and lack of information on T-cell reactivity for potential cross-reactivity in some small subset of patients."
- *Dr. Kirkpatrick:* "I am concerned about the post hoc analysis that had to be done to yield a positive result; I continue to have concerns about bone…statistics and the bias issue, and I still have the concern over the very rare incidence/potential of a drug having a relatively low incidence, but catastrophic, event occur."
- *Dr. MacLaughlin:* "I have concerns about safety issues going forward re pregnancy. I have to rely on statistical and clinical correlations. And I'm unconvinced of its effectiveness. Also (I'm concerned about) some of the biases coming up in the statistical arguments. There is a lot of promise here, and it's important to try to use recombinant materials to replace surgical procedures. Post hoc analysis…is always a little flag to me, and I had trouble getting past that."
- *Dr. Propert:* "I had concerns about the efficacy analysis and conclusions…There is inadequate data to assess immunological safety at this time."
- *Dr. McCormick:* "I don't think that the efficacy was convincingly demonstrated. Still, I think it's a safe product as much as we can tell at this state. For some patients I thought it was effective. It did create bone, and I think it would have been a nice tool to have. I would not have approved it without significant indications along with it, but that is moot at this point."

The panel chair then asked the panel members what is needed to make the PMA approvable. They responded:

- *Dr. McCormick, a neurosurgeon:* "A new study (is needed). The data are what they are. It may be reasonable to repeat the study under more realistic circumstances or contemporary circumstances…My biggest concern here was such a narrow population. That's not the population you can show its efficacy in. The presence of any bone on CT was a real problem."
- *Dr. Propert, a statistician:* "A new study (is needed), correcting some of the flaws of this, some of which I realize were historical, and putting the CT scans right up front."
- *Dr. MacLaughlin, a biochemist:* "It's possible to allay some of the immune issues in animal models."
- *Dr. Kirkpatrick:* "Trying to do something on the immune memory that I can understand would be helpful. A clinical study, using contemporary (approaches) so you don't have dropouts. Looking at instrument fusion with OP-1 vs. autograft would be a reasonable straightforward study, looking at bridging bone with a CT scan."
- *Dr. Jason, an immunologist:* "The study would clearly have to involve a different patient population with different control procedure. Do some cellular assays *in vitro* and ideally as people get enrolled, look at cellular reactivity prior to the implant and then look at function after the implant. Break it down by people who have natural antibodies compared to people who make antibodies after the procedure. You'd get some sense on whether you have reason to worry. You say something like 3% ends up in the bloodstream, but to characterize that and see how much of that is basically the profile – whether it's aggregate or not would be very useful."
- *Dr. Rao, a spine surgeon:* "I can't offer anything in terms of suggestions that they haven't thought of already. However, I will defer to them on the question of any new studies. On the data already available on the CT scans… look at unilateral or bilateral...assess if there's any way to look at something other than the presence of bone. Also, I'd like to see the area of antibodies across the maternal/fetal membranes."
- *FDA's Melkerson:* "Typically in orthopedics we look at 24 months. With a bone-forming agent, shorter-term studies are something that need to be a year based on the data we've seen."
- *Dr. Kirkpatrick:* "It depends on if the FDA is concerned about the subset of bridging bone. That's something that could be seen in a year. In contemporary practice with the newer techniques on CT generally we can see confluent bone…if the FDA's question is related solely to fusion and bridging bone that could be done in a year."

♦