Food and Drug Administration

Executive Summary for P050036
Medtronic's AMPLIFY™ rhBMP-2 Matrix

Orthopaedic and Rehabilitation Devices Advisory Panel

July 27, 2010

TABLE OF CONTENTS

Section	Page Number
INTRODUCTION	**4**
Rationale for Presentation to the Panel	**4**
FDA Questions to the Panel	**4**
BACKGROUND INFORMATION	**5**
Applicant Name and Address	**5**
Indications for Use	**5**
Contraindications	**5**
Warnings	**6**
Device Description	**7**
Recombinant Human Bone Morphogenetic Protein-2 (rhBMP-2) Solution	8
Compression Resistant Matrix (CRM)	8
Posterior Spinal Fixation System	8
Overview of Product Preparation	9
Total Dose and Concentration	9
Regulatory History	**9**
NON-CLINICAL DATA	**10**
Reproductive Toxicity	**10**
Study Design	10
Antibody Analysis	11
Skeletal Analysis	11
Device Labeling	12
CLINICAL STUDY DESCRIPTION	**14**
Purpose	**14**
Study Design	**14**
Primary Study Endpoint / Success Criteria	**15**
Secondary Study Endpoints	**15**
Inclusion/Exclusion Criteria	**16**
Evaluations	**18**
Adverse Events	**18**
Statistical Analysis Plan	**20**
Randomization and Blinding	20
Hypotheses Tested	20
Statistical Methodology	20
CLINICAL STUDY RESULTS	**22**
Patient Accounting	**22**
Demographics and Preoperative Characteristics	**23**
Surgical and Discharge Information	**25**
SAFETY EVALUATION	**26**
All Adverse Events	**26**
Serious Adverse Events	**31**

Cancer Events **33**
Detailed Information on Specific Adverse Event Categories **38**
Patient Deaths 38
Antibody Monitoring 41
All Device-Related Adverse Events 42
Serious Device-Related Adverse Events 43
Secondary Surgical Interventions **44**
Safety Evaluation Summary **45**
EFFECTIVENESS EVALUATION **46**
Primary Endpoint (Overall Clinical Success) **46**
Per-Protocol Results 48
Missing Equals Failure Results 48
Selected Components of Overall Success **50**
Oswestry Disability Index (ODI) 50
Neurological Status 50
Fusion Status 51
Secondary Effectiveness Endpoints **52**
Leg Pain Numerical Rating Scale (NRS) 52
Back Pain Numerical Rating Scale (NRS) 52
Short Form 36 (SF-36) 53
Clinical Study Discussion **55**
ADDITIONAL CANCER DATA **56**
Evaluation of Cancer in Clinical Studies **56**
Overall Results 58
FDA Comments on Evaluation of Cancer in Clinical Studies 60
Medicare Study of Pancreatic Cancer **61**
Objective 61
Enrollment and follow-up 62
Outcomes 62
Statistical Plan 62
Validation 63
Descriptive Results 63
Analytic Results 63
Validation Results 64
FDA Comments on Medicare Study of Pancreatic Cancer 65
POST-APPROVAL STUDY **67**
Objective **67**
Enrollment and Follow-up **67**
Outcomes **68**
Statistical Plan **68**
Validation **68**
FDA Comments on Proposed Post-Approval Study **68**

INTRODUCTION

The subject of this Executive Summary is Medtronic's AMPLIFY™ rhBMP-2 Matrix premarket approval (PMA) application, P050036. The AMPLIFY™ rhBMP-2 Matrix is a combination product consisting of devices (Compression Resistant Matrix or CRM, as well as metallic posterior spinal instrumentation) and a drug (recombinant human Bone Morphogenetic Protein 2 or rhBMP-2). This application has been reviewed by staff in the Center for Devices and Radiological Health (CDRH) and Center for Drug Evaluation and Research (CDER) of the Food and Drug Administration (FDA). Your time and effort in the review of this application are greatly appreciated.

Rationale for Presentation to the Panel

The FDA presents this PMA to the Orthopaedic and Rehabilitation Devices Advisory Panel for the following reasons:

- The AMPLIFY™ rhBMP-2 Matrix is the first combination product utilizing rhBMP-2 for posterolateral fusion treatment of degenerative disc disease (DDD).

- The number of cancer events reported in the IDE clinical study of AMPLIFY™ rhBMP-2 Matrix is higher in patients treated with the investigational device compared to the number of cancer events in patients treated with the control device.

- The number of cancer events reported in all clinical trials of rhBMP-2 may be higher than the number of cancer events reported in the respective comparison groups.

FDA Questions to the Panel

The FDA would like to the Panel to provide responses to several questions regarding the safety and effectiveness data contained in P050036. These questions are located in the "FDA Panel Questions" section of the Panel package.

BACKGROUND INFORMATION

Applicant Name and Address

Medtronic Sofamor Danek USA, Inc.
1800 Pyramid Place
Memphis, Tennessee 38132

Indications for Use

The following Indications for Use were specified in the IDE clinical study for the AMPLIFY™ rhBMP-2 Matrix System:

> "The device is to be used for posterolateral spinal fusion treatment of single level lumbar (L1-S1) degenerative disc disease (DDD) in skeletally mature patients who failed at least 6 months of nonoperative conservative treatment. The device is to be used in conjunction with the CD HORIZON® Spinal System, a previously cleared temporary posterior supplemental fixation device."

The following Indications for Use are proposed by the sponsor in the PMA:

> "The AMPLIFY™ rhBMP-2 Matrix is indicated as an alternative to autogenous bone graft for spinal fusion procedures in skeletally mature patients with degenerative disc disease (DDD) at one level from L1-S1. DDD is defined as discogenic back pain with degeneration of the disc confirmed by patient history and radiographic studies. DDD patients may also have up to Grade 1 spondylolisthesis or retrolisthesis at the involved level. Patients receiving AMPLIFY™ rhBMP-2 Matrix should have had at least 6 months of non-operative treatment prior to implantation of AMPLIFY™ rhBMP-2 Matrix. AMPLIFY™ rhBMP-2 Matrix is to be implanted via a posterolateral approach and must be used in conjunction with a metallic posterior supplemental fixation device that is indicated for temporary stabilization of the spine."

Contraindications

The sponsor proposes that use of the AMPLIFY™ rhBMP-2 Matrix be contraindicated in the following cases:

- patients with a known hypersensitivity to recombinant human Bone Morphogenetic Protein-2 (rhBMP-2), bovine Type I collagen, or other components of the formulation

- in the vicinity of a resected or extant tumor, in patients with any active malignancy, or in patients undergoing treatment for a malignancy

- patients who are skeletally immature (<18 years of age or no radiographic evidence of epiphyseal closure)

- pregnant women (the potential effects of rhBMP-2 on the human fetus have not been evaluated)

- patients with an active infection at the operative site or with an allergy to one of the metals used in the posterior supplemental fixation device (such as titanium, stainless steel, or cobalt chromium alloy).

Warnings

The sponsor proposes that the following warnings be included in the labeling for the AMPLIFY™ rhBMP-2 Matrix, which is based on the labeling for other PMA-approved products containing rhBMP-2 (P000054, P000058, and P050053):

- In an experimental rabbit study, rhBMP-2 has been shown to elicit antibodies that are capable of crossing the placenta. Reduced ossification of the frontal and parietal bones of the skull was noted infrequently (<3%) in fetuses of rabbit dams immunized to rhBMP-2; however, there was no effect noted in limb bud development. There are no adequate and well-controlled studies in human pregnant women. Women of childbearing potential should be warned by their surgeon of potential risk to a fetus and informed of other possible orthopedic treatments.

- Women of childbearing potential should be advised that antibody formation to rhBMP-2 or its influence on human fetal development has not been completely assessed. In the clinical trial supporting the safety and effectiveness of AMPLIFY™ rhBMP-2 Matrix, 15/234 (6.4%) patients treated with AMPLIFY™ rhBMP-2 Matrix and 5/217 (2.3%) patients treated with autograft bone developed antibodies to rhBMP-2, based on an anti-human immunoglobulin antibody-based ELISA. The effect of maternal antibodies to rhBMP-2, as might be present for several months following device implantation, on the unborn fetus is unknown. Additionally, it is unknown whether fetal expression of BMP-2 could re-expose mothers who were previously antibody positive. Theoretically, re-exposure may elicit a more powerful immune response to BMP-2 with possible adverse consequences for the fetus. However, pregnancy did not lead to an increase in antibodies in the rabbit study. Studies in genetically altered mice indicate that BMP-2 is critical to fetal development and that a lack of BMP-2 activity may

cause neonatal death or birth defects. It is not known if anti-BMP-2 antibodies may affect fetal development or the extent to which these antibodies may reduce BMP-2 activity.

- AMPLIFY™ rhBMP-2 Matrix should not be used immediately prior to or during pregnancy. Women of childbearing potential should be advised not to become pregnant for one year following treatment with AMPLIFY™ rhBMP-2 Matrix.

- The safety and effectiveness of AMPLIFY™ rhBMP-2 Matrix in nursing mothers has not been established. It is not known if BMP-2 is excreted in human milk.

- The safety and effectiveness of AMPLIFY™ rhBMP-2 Matrix with other spinal implants, implanted at locations other than the lower lumbar spine, or used in surgical techniques other than a posterolateral technique have not been established.

- Inappropriate use of the product, such as preparing it differently than prescribed or compressing the rhBMP-2/CRM construct more than necessary, may change the concentration of the rhBMP-2, which may cause complications.

Device Description

The AMPLIFY™ rhBMP-2 Matrix is a combination product consisting of rhBMP-2 solution, a collagen compression-resistant matrix (CRM), and metallic posterior spinal instrumentation. The device is used for posterolateral fusion treatment of single level lumbar (L1-S1) degenerative disc disease (DDD).

The AMPLIFY™ rhBMP-2 Matrix would be marketed as a single 40mg (20cc) kit, composed of the following:

(2) vials of sterile rhBMP-2 (20mg each)
(2) packages of two (2) sterile 5cc Compression Resistant Matrices (CRMs)
(2) 10mL vials of sterile water
(2) 10mL sterile syringes with 20G 1½" needles
(4) 3mL sterile syringes with 20G 1½" needles

Recombinant Human Bone Morphogenetic Protein-2 (rhBMP-2) Solution

Bone morphogenetic proteins are members of the transforming growth factor-beta (TGF-β) superfamily. Human bone morphogenetic protein-2 (hBMP-2) is osteoinductive, inducing bone formation in vivo. The primary mode of action for hBMP-2 is to differentiate mesenchymal cells derived from the periosteum (soft tissues surrounding the implant site) or the bone marrow stroma into cartilage and subsequently bone cells, resulting in endochondral or intramembranous bone formation.

Recombinant human bone morphogenetic protein-2 (rhBMP-2), also known as dibotermin alfa, is secreted from genetically engineered cultures of Chinese hamster ovary (CHO) cells. The active rhBMP-2 protein molecule consists of a disulfide-linked dimer with two major subunit species of 114 and 131 amino acids, respectively.

The rhBMP-2 is intended to stimulate bone formation during the posterolateral fusion process. The rhBMP-2 is provided in two vials as a lyophilized powder. At the time of surgery, the rhBMP-2 is reconstituted using the provided sterile water and injected onto the CRMs.

Please note the rhBMP-2 protein used in the AMPLIFY™ rhBMP-2 Matrix is chemically identical to that used in the INFUSE® Bone Graft component of three PMA-approved products (P000054, P000058, and P050053).

Compression Resistant Matrix (CRM)

The CRM provides the matrix for delivery of the rhBMP-2 and is used to maintain the rhBMP-2 at the site of implantation. The CRM also serves as a scaffold on which new bone may form. The CRM consists of an absorbable bovine Type I collagen with embedded biphasic calcium phosphate granules (15% hydroxyapatite (HA) and 85% tricalcium phosphate (TCP)). Each CRM is 4.67cm long, 0.95cm wide, and 1.13cm high. The granules range in diameter from 0.5mm to 1.6mm.

Please note the CRM materials used in the AMPLIFY™ rhBMP-2 Matrix are identical to the 510(k)-cleared MASTERGRAFT® Matrix (K023553). The collagen portion of the CRM is identical to that used in two other approved PMAs (P000054, P000058, and P050053).

Posterior Spinal Fixation System

The posterior fixation system provides stabilization of the affected spinal segments in order to facilitate fusion. The rhBMP-2 solution and CRM must be used in conjunction with a metallic posterior spinal fixation system cleared for treatment of DDD.

In the IDE clinical study, the posterior fixation provided was titanium-only components of the 510(k)-cleared CD HORIZON® Spinal System. However, the AMPLIFY™ rhBMP-2

Matrix will not be sold with posterior instrumentation; a posterior fixation system must be procured separately.

Overview of Product Preparation
At the time of surgery, 5.4 mL of sterile water are drawn using one of the 10mL syringes and then injected into one of the 20mg vials of lyophilized rhBMP-2 for reconstitution. A 3mL syringe is used to draw 2.5mL of the reconstituted rhBMP-2 solution, which is then distributed uniformly on one of the 5cc CRMs. A second 3mL syringe is used to wet a second 5cc CRM. These steps are repeated using a new of set of rhBMP-2/sterile water vials and 5cc CRMs. Once prepared, two (2) wetted CRMs are placed across the decorticated transverse processes on each side of the level to be treated. The placement of the posterior fixation system at the treated level (i.e., before or after implantation of AMPLIFY™ rhBMP-2 Matrix) is left to the discretion of the surgeon.

Total Dose and Concentration
Once prepared according to the instructions for use and implanted, the AMPLIFY™ rhBMP-2 Matrix delivers a total of ~37.0mg of rhBMP-2, with a concentration of ~2.0mg/cc of CRM.

Regulatory History

The AMPLIFY™ rhBMP-2 Matrix is not marketed in the United States or any foreign country, and has only been used in IDE studies in the United States.

As previously stated, the rhBMP-2 and Compression Resistant Matrix (CRM) components of the AMPLIFY™ rhBMP-2 Matrix have been marketed in the United States as part of other PMA-approved and 510(k)-cleared devices. The rhBMP-2 is marketed in the United States as part of the INFUSE® Bone Graft component for three PMA-approved devices:

- P000054 – INFUSE® Bone Graft *(treatment for acute, open tibial shaft fractures)*
- P000058 – INFUSE® Bone Graft/LT-CAGE Lumbar Tapered Fusion Device *(anterior lumbar interbody fusion (ALIF) treatment of single-level lumbar (L2-S1) DDD)*
- P050053 – INFUSE® Bone Graft *(alternative to autogenous bone graft for sinus augmentations, and for localized alveolar ridge augmentations for defects associated with extraction sockets)*

The Compression Resistant Matrix (CRM) component has been marketed in the United States as a bone void filler, including for use in posterolateral fusion procedures, since 2003 (K023553 – MASTERGRAFT® Matrix).

NON-CLINICAL DATA

As previously stated, the rhBMP-2 and CRM components have been individually approved via PMA application, e.g., rhBMP-2, or cleared for market via 510(k) submission, e.g., CRM, in a number of applications. In addition, extensive pre-clinical data have been provided as part of P000054, P000058, and P050053 to characterize and to support the safety and effectiveness of rhBMP-2.

Some pre-clinical studies were conducted to evaluate the specific rhBMP-2/CRM combination used in the AMPLIFY™ rhBMP-2 Matrix. These studies are summarized in the "Proposed Summary of Safety and Effectiveness Document" section of the Panel package and are listed here.

- CRM Development
- CRM Carrier Efficacy Studies
- Selection of rhBMP-2 Concentration and Ceramic Composition

At this time, the FDA has no comments or concerns regarding any of the non-clinical data presented, except for the following issue.

Reproductive Toxicity

In the approval orders for PMAs P000054 and P000058, the sponsor was asked to conduct studies to evaluate the potential for an immune response to rhBMP-2 to interfere in embryonic development in rabbits. In response, the sponsor provided a reproductive toxicology study conducted by Wyeth, entitled, "WAY-167639 (rhBMP-2): Intramuscular Immunization Developmental Toxicity Study in Rabbits (Protocol 03_0357)."

The purpose of this study was to determine whether the immune response to rhBMP-2 could influence offspring development (e.g., from embryo/fetal mortality to congenital abnormalities of non-lethal severity). The study was conducted in accordance with ICH guidance, i.e., *Guideline on detection of toxicity to reproduction for medicinal products*, and GLP regulations.

Study Design
A total of 50 female New Zealand White (NZW) rabbits were divided into experimental (30 animals) and control (20 animals) groups. The experimental animals were treated with a solution containing Titermax Gold® and rhMBP-2 (2mg rhBMP-2 resulting in a final rhBMP-2 concentration of 2.55mg/mL) in a 1:1 ratio. The control animals received

a 1:1 mixture of Titermax Gold® and sterile saline. Each animal received intramuscular injections of the assigned treatment on days 1, 8, 22, and 43, with the last injection occurring 3 days prior to mating. Antibody titer was assessed prior to mating and the animals with the greatest titer were allowed to mate first. All animals were sacrificed 29 days after mating and the fetuses were examined for abnormalities. Evaluations consisted of the mortality, clinical observations, abortion rate, body weight, food consumption, gravid uterine weight, hysterotomy findings on gestation day (GD) 29 (corpora lutea, litter size, embryo/fetal mortality), and postmortem observations, fetal sex, weight, and external, palatal, visceral, and skeletal anomalies, and placental appearance. Serum titers of anti-rhBMP-2 antibodies and rhBMP-2 neutralizing antibody activity were also evaluated.

The following sections highlight some key findings from this study.

Antibody Analysis
Pregnancy did not increase the immune response of the mothers. Antibody levels in the fetuses were similar to those of the mothers. This suggests that anti-rhBMP-2 antibodies produced by the mothers were capable of crossing the placenta and remaining at a level similar to those of the mother. For mothers that had detectable levels of neutralizing anti-rhBMP-2 antibodies, these were also found in the fetuses, again indicating that they were able to cross the placenta.

Skeletal Analysis
Fetal skeletal observations included the following:

- Reduced ossification of frontal bones
- Reduced ossification of parietal bones
- Reduced ossification of metacarpal bones
- Reduced number of cervical ribs
- Misaligned caudal vertebrae
- Bipartite sternebrae and vertebral centrum
- Reduced ossification of pelvic girdle and hyoid
- Unossified talus

In the experimental group, the number of observed reductions in ossification of the frontal and parietal bones was considered to be increased compared to historical controls whereas other observations, e.g., ossification of the metacarpal bones, reduced number of cervical ribs, etc…, were within historical control data and comparable to the control cohort of the study. The sponsor believes that the observations of reduced ossification of the frontal and parietal bones are of no toxicological significance since reduced ossification during gestation is believed to be of a reversible nature. The sponsor cites 4 published papers to support their contention. In particular, they cite a

1970 publication entitled, *Ossification of the rat and mouse skeleton in the perinatal period* as supporting information. The sponsor indicates that there was no apparent correlation between fetal antibody titer levels and incidence of multiple sites of reduced ossification nor was there a correlation between the presence of neutralizing activity and reduced ossification – 50% of fetuses exhibiting neutralizing antibody activity displayed normal skeletal ossification.

The sponsor's interpretations and the FDA's interpretations are different. In summary, the FDA considers that reduced (or delayed) ossification possibly as a result of exposure to anti-rhBMP-2 antibodies may not be reversible, for several reasons.

First, research into bone inducing factors such as BMP had only initiated during the late 60's and molecular identification of individual BMP members was not accomplished until the 80's and later. It is clear that an observation such as this, i.e., reduced ossification in fetal growth in association with antibodies to mediators of bone development (BMPs), would not have been previously considered. Therefore, the reversible nature of a reduction in ossification due to anti-BMP-2 antibodies is not assured, based solely on the 4 publications cited, and this observation should be carefully considered.

Second, the sponsor indicates that there was no apparent correlation between fetal antibody titer levels and incidence of multiple sites of reduced ossification nor was there a correlation between the presence of neutralizing activity and reduced. However, a lack of power in the numbers evaluated in the study, or variation in epitope targeting of the antibodies, could possibly explain any apparent lack of correlations.

Finally, the fetuses in the study were not allowed to develop and mature. As a result, it is not known if the delayed/reduced ossification would resolve at some point after birth or would remain. Because there was no apparent correlation between fetal antibody titer levels and incidences of reduced ossification, nor was there a correlation between the presence of neutralizing activity and reduced ossification(i.e., ~50% of fetuses exhibiting neutralizing antibody activity displayed normal skeletal ossification), the sponsor would contend that the observations were incidental to, and not caused by, anti-rhBMP-2 antibodies. However, the FDA believes there is an inadequate sample size to support this conclusion.

Because of these confounding factors, the FDA believes an additional study should be conducted to evaluate further how the immune response to rhBMP-2 affects offspring development, with longer-term evaluation of fetuses from animals treated with rhBMP-2.

Device Labeling

The FDA and the sponsor were unable to reach agreement on the interpretation of the results of the above study. To address our concerns, the sponsor included the following

language as black box warnings in the device labeling for P000054, P000058, and P050053 and is also proposed for the subject device:

- In an experimental rabbit study, rhBMP-2 has been shown to elicit antibodies that are capable of crossing the placenta. Reduced ossification of the frontal and parietal bones of the skull was noted infrequently (<3%) in fetuses of rabbit dams immunized to rhBMP-2; however, there was no effect noted in limb bud development. There are no adequate and well-controlled studies in human pregnant women. Women of childbearing potential should be warned by their surgeon of potential risk to a fetus and informed of other possible orthopedic treatments.

- Women of childbearing potential should be advised that antibody formation to rhBMP-2 or its influence on human fetal development has not been completely assessed. In the clinical trial supporting the safety and effectiveness of AMPLIFY™ rhBMP-2 Matrix, 15/234 (6.4%) patients treated with AMPLIFY™ rhBMP-2 Matrix and 5/217 (2.3%) patients treated with autograft bone developed antibodies to rhBMP-2, based on an anti-human immunoglobulin antibody-based ELISA. The effect of maternal antibodies to rhBMP-2, as might be present for several months following device implantation, on the unborn fetus is unknown. Additionally, it is unknown whether fetal expression of BMP-2 could re-expose mothers who were previously antibody positive. Theoretically, re-exposure may elicit a more powerful immune response to BMP-2 with possible adverse consequences for the fetus. However, pregnancy did not lead to an increase in antibodies in the rabbit study. Studies in genetically altered mice indicate that BMP-2 is critical to fetal development and that a lack of BMP-2 activity may cause neonatal death or birth defects. It is not known if anti-BMP-2 antibodies may affect fetal development or the extent to which these antibodies may reduce BMP-2 activity.

- AMPLIFY™ rhBMP-2 Matrix should not be used immediately prior to or during pregnancy. Women of childbearing potential should be advised not to become pregnant for one year following treatment with AMPLIFY™ rhBMP-2 Matrix.

- The safety and effectiveness of AMPLIFY™ rhBMP-2 Matrix in nursing mothers has not been established. It is not known if BMP-2 is excreted in human milk.

The FDA will be asking the Panel to comment on the adequacy of the non-clinical data presented to date, as well as the proposed device labeling, in addressing the potential risks to health for women of childbearing potential, women who are pregnant, and the human fetus.

CLINICAL STUDY DESCRIPTION

This section summarizes the IDE clinical study protocol. For a complete investigational protocol, please refer to Attachment A.

Purpose

The purpose of this clinical study, as stated in the IDE investigational plan, was to demonstrate the safety and effectiveness of the AMPLIFY™ rhBMP-2 Matrix System for patients requiring 1-level posterior fusion procedures of the lumbar, and/or sacral spine (L1-S1) for degenerative disc disease. The ability for this device to provide successful fusion of spinal segments, while positively affecting clinical outcomes, was assessed and compared to bilateral posterolateral implantation of autogenous bone harvested from the iliac crest and stabilized, as well, with the CD Horizon posterior supplemental fixation device.

The primary safety objectives were assessed by evaluating all patients for serious adverse events that were device- or device/surgical procedure-related, additional surgical interventions classified as failures through 24 months, neurological success based on the results from neurological evaluations at 24 months, and production of antibodies to rhBMP-2 and bovine Type 1 collagen.

The primary efficacy objectives were assessed by evaluating functional success based on results from the Oswestry Disability Index at 24 months, as well as radiographic evaluation of fusion success.

Study Design

The sponsor provided data from the pivotal multi-center, prospective, randomized, concurrently controlled, non-blinded, non-inferiority trial of the AMPLIFY™ rhBMP-2 Matrix System compared to posterolateral spinal fusion using autogenous bone in patients with degenerative disc disease as defined in the study inclusion/exclusion criteria outlined below.

The original study design was approved for 480 patients at 30 sites (240 investigational, 240 control). 463 patients were randomized and implanted at 29 clinical centers in the United States for a total of 239 investigational and 224 control patients. The first patient was enrolled on March 29, 2002, and the last was enrolled on March 26, 2004. The most recent date of database closure for analyses was February 3, 2009.

Primary Study Endpoint / Success Criteria

Individual patient success (i.e., overall success) was determined at 24 months and was defined as a composite endpoint. A patient was considered a success if all of the following criteria regarding safety and efficacy were met:

- Fusion defined as
  - Evidence of bridging trabecular bone (continuous bony connection from superior transverse process to the inferior transverse process on both sides, based on radiographs or CT scans.
  - No evidence of motion
    - no more than 3 mm difference in translation on lateral flexion/extension radiographs
    - less than 5° difference in angular motion between flexion and extension)
  - The absence of radiolucent lines through the fusion mass

  Fusion outcomes were used to indicate radiographic success
- Function defined as
  - Improvement of at least 15 points on the Oswestry Disability Index (ODI, version 2) graded on a 100 point scale as compared to baseline
- Maintenance or improvement in neurological status with no new permanent neurological deficits as compared to baseline
- Absence of "device associated or device/surgical procedure associated" adverse events
- Freedom from additional surgical intervention defined as revision, non-elective removal or supplemental fixation/fusion at the affected level

Secondary Study Endpoints

The secondary study endpoints were determined at 24 months and evaluated to compare the success rates of the individual safety and effectiveness endpoints, including operative measurements. Secondary endpoints included:

- Back and Leg Pain
  - For each component, Preop score-Postop score≥0 where score=pain intensity + pain duration
- Hip (Donor Site) Pain
  - Pain level measured in the control group
  - To support a claim of benefit for the investigational device
- General Health Status
  - SF-36 scores split into their MCS and PCS components and scored separately

  - Success will be defined as maintenance or improvement of the individual component scores
    - PCS_{postop}-PCS_{preop}≥0
    - MCS_{postop}-MCS_{preop}≥0
- Global perceived effect
- Patient Satisfaction
- Doctor's Perception of Results
- Work status
- Antibody levels
  - Subjects with authentic elevated antibody levels will be compared to those without elevated levels

The study was considered a success if the Overall Success Rate for the device was determined to be non-inferior as compared to the Overall Success Rate for the control. As per the protocol, neither investigators nor patients were blinded, primarily because of the second surgical site necessary for collection of the iliac crest grafts. The radiologists were blinded as to the treatment received. Bayesian statistical methods are used to assess non-inferiority and superiority hypotheses for outcome variables between the two groups with respect to:

- Study hypotheses including overall success, individual effectiveness variables, and neurological status
- Cumulative event rates for adverse events and additional surgical interventions
- Comparative means for operative time, blood loss, and number of hospital days

Inclusion/Exclusion Criteria

Table 1 – Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
• Degenerative disc disease (DDD) accompanied by back pain of discogenic origin, with or without leg pain, with degeneration of the disc confirmed by patient history and radiographic studies • DDD present if one or more of the following was present: • instability (defined as angulation ≥ 5° and/or translation ≥ 4 mm, based on flexion/extension radiographs • osteophyte formation • decreased disc height • thickening of ligamentous tissue • disc degeneration or herniation and/or • facet joint degeneration • Requires fusion of a single level disc space	• Primary diagnosis of a spinal disorder other than degenerative disc disease with Grade 1 or less spondylolisthesis at the involved level • Previous lumbar spinal fusion procedure at the involved level • Required spinal fusion at more than one lumbar level • Conditions that might have been associated with diagnosis of osteoporosis including: – Postmenopausal Non-Black female over 60 years of age and weighing less than 140 pounds. – Postmenopausal female that had sustained a nontraumatic hip, spine, or wrist fracture. – Male over the age of 70. – Male over the age of 60 that had sustained a

Table 1 – Inclusion and Exclusion Criteria	
Inclusion Criteria	**Exclusion Criteria**
from L1 to S1 • Preoperative Oswestry Disability Index score of ≥ 30 • No greater than Grade 1 spondylolisthesis utilizing Meyerding's Classification • At least 18 years of age, inclusive, at the time of surgery • Non-responsive to non-operative treatment (e.g., bed rest, physical therapy, medications, spinal injections, manipulation, and/or TENS) for a period of 6 months • If of child-bearing potential, non-pregnant, non-nursing, and agreed to use adequate contraception for 1 year following surgery • Willing and able to comply with study plan and able to understand and sign Patient Informed Consent Form	non-traumatic hip or spine fracture. If the level of BMD is a T score of -3.5 or a T score of -2.5 with vertebral crush fracture, the patient was excluded from the study • Previous diagnosis of osteopenia or osteomalacia • Conditions that require postoperative medications that interfere with fusion, such as steroids or prolonged use of non-steroidal anti-inflammatory drugs, excluding routine perioperative non-steroidal anti-inflammatory drugs. This does not include low dose aspirin for prophylactic anticoagulation. • Overt or active bacterial infection, either local or systemic • Presence of active malignancy or prior history of malignancy (except for basal cell carcinoma of the skin • Documented titanium allergy or intolerance • Mentally incompetent • 'Waddell Signs of Inorganic Behavior' score of 3 or greater • Prisoner • Alcohol and/or drug abuser currently undergoing treatment for alcohol and/or drug abuse • Received drugs that may interfere with bone metabolism within two weeks prior to the planned date of spinal surgery (e.g., steroids or methotrexate); • History of autoimmune disease (e.g., SLE or dermatomyositis) • History of exposure to injectable collagen or silicone implants • History of hypersensitivity to protein pharmaceuticals (monoclonal antibodies or gamma globulins) or collagen • Received treatment with an investigational therapy (device and/or pharmaceutical) within 28 days prior to implantation surgery or such treatment is planned during the 16 weeks following rhBMP-2/CRM implantation • Received any previous exposure to any/all BMPs of either human or animal extraction • History of allergy to bovine products or a history of anaphylaxis • History of any endocrine or metabolic disorder known to affect osteogenesis

Evaluations

Patients were evaluated preoperatively (within 6 months of surgery), intra-operatively, and postoperatively at 6 weeks, 3, 6, 12, and 24 months, and then at 36, 48, and 60 months, as per the IDE protocol requiring 5 year follow-up. Fusion was evaluated at 6, 12, and 24 months. Complications and adverse events were evaluated over the course of the clinical trial. The protocol also included measurements of antibodies to rhBMP-2 and bovine Type I collagen screening in both investigational and control patients pre-operatively and at 6 weeks, 3, 6, and 12 months. At each evaluation time-point, the primary and secondary clinical and radiographic outcome parameters were assessed. Success was determined from data collected during the initial 24 months of follow-up.

A summary of all study evaluations is provided in Table 2.

Adverse Events

An adverse event was defined as "…any clinically adverse sign, symptom, syndrome, or illness that occurs or worsens during the operative and postoperative periods of the trial, regardless of causality, that is not being otherwise measured in the trial." Pain, neurological and functional symptoms were considered adverse events when a patient's complaint for any of these symptoms resulted in an unscheduled visit or when a patient presented with new or worsening pain, neurological and/or functional symptoms as compared to the previous visit. All adverse events reported in the Clinical Summary Report for 24-month data submitted on April 3, 2009 underwent review by an independent Data Safety Monitoring Board (DSMB) with respect to the severity (complication or observation) and the relatedness of each event. Final classification of adverse events was determined by the sponsor.

Table 2 – Schedule of Study Assessments							
	Pre/Perioperative		**Postoperative**				
Procedure	**Preoperative**	**Surgery/Hospital Discharge**	**6wks ±2wks**	**3mo ±2wks**	**6mo ±1mo**	**12mo ±2mo**	**24mo ±2mo & annually**
Preoperative Information							
Confirm Patient Eligibility	X						
Obtain Informed Consent	X						
Obtain HIPAA Authorization	X						
Case Report Forms							
Patient Enrollment	X						
Patient Qualification Form	X						
Preoperative Data	X						
Preoperative Patient Survey	X						
Preoperative Oswestry	X						
Blood Specimen for Antibody Monitoring	X		X	X	X	X	
SF-36 Questionnaire	X		X	X	X	X	X
Neurological Status	X	X	X	X	X	X	X
Surgery Data		X					
Hospital Discharge		X					
Postoperative Data			X	X	X	X	X
Postoperative Patient Survey			X	X	X	X	X
Postoperative Oswestry			X	X	X	X	X
Back & Leg Pain Questionnaire	X		X	X	X	X	X
Hip Pain Questionnaire (Control Group Only)		X	X	X	X	X	X
Radiographic Data	X	X	X	X	X	X	X
Adverse Event (if any)		X	X	X	X	X	X
Outstanding AE (if applicable)			X	X	X	X	X
Study Deviation (if applicable)	X	X	X	X	X	X	X
Patient Accountability			X	X	X	X	X
Radiographic Procedures							
Anterior/Posterior*	X	X	X	X	X	X	X
A/P Ferguson (if L5-S1)**	X	X	X	X	X	X	X
Lateral	X	X	X	X	X	X	X
Lateral Flexion/Extension	X			X	X	X	X
CT and /or MRI	X						
CT (high resolution axial, sagittal and coronal reformatting)		X			X	X	X
DEXA Scan***	X						

*If L1-2, L2-3, L3-4, L4-5 **If L5-S1 *** If any part of exclusion criteria regarding previous diagnosis of osteoporosis was marked yes..

For secondary surgical interventions, the protocol specified that a revision, supplemental fixation or non-elective implant removal would be classified as a treatment "failures." A reoperation, elective removal, or other surgical procedures were not classified as a treatment "failures." All patients who underwent secondary surgical procedures were followed for the duration of the study but data was censored at time of failure.

Statistical Analysis Plan

Randomization and Blinding

The randomization was in a 1:1 ratio of investigational to control patients. However, at different sites, there were markedly uneven randomization with respect to investigator and what procedure was performed. For example, at one site, one surgeon performed 9 investigational and 1 control procedures and another investigator performed 5 control procedures and no investigational ones.

As per the protocol, neither investigators nor patients were blinded, primarily because of the second surgical site necessary for collection of the iliac crest grafts. The lack of patient and surgeon blinding in this study is an important limitation of the study design and an unaddressed source of bias. The lack of blinding could potentially have led to reporting bias among patients and investigators, in favor of the investigational device or against the control. This may be particularly problematic for subjective assessments such as patient reported outcomes. For radiographic measures, the radiologists were blinded as to the treatment received.

Hypotheses Tested

This was a non-inferiority trial and a fixed non-inferiority margin of 10% was required by FDA. The primary objective was to demonstrate non-inferiority in the primary endpoint (patient success) at 24 months after surgery. Secondary objectives were to demonstrate superiority in the primary endpoint of the investigational group over the control group at 24 months after surgery and to demonstrate non-inferiority (and possibly superiority) in the individual effectiveness variables (radiographic fusion and Oswestry pain/disability status), as well as maintenance or improvement in neurological status. According to the data presented the sponsor achieved their primary endpoint – non-inferiority in overall success, as well as non-inferiority in the secondary endpoints of ODI, neurological outcome, and fusion. Superiority was established only in fusion.

Statistical Methodology

The sponsor utilized three major statistical plans for different data as follows:

- For demographic and preoperative information
    - descriptive analyses of proportions (qualitative variables)
    - means and standard deviations (quantitative variables)
- For pooling data across sites

- Breslow-Day test to assess homogeneity
  - primary endpoints
    - Oswestry
    - neurological
    - fusion
  - overall success at 24 months

- Non-inferiority and superiority hypotheses for outcome variables
  - Bayesian statistical methods
    - outcome variables
      - overall success
      - individual effectiveness variables
      - neurological status
    - cumulative event rates
      - adverse events
      - additional surgical interventions
  - comparative means
    - operative time
    - blood loss
    - number of hospital days

Three different analysis datasets (primary, per-protocol, and missing-equals-failure datasets) were defined. The PMA is based on the final analysis of the data evaluated up to 24-month visits from the whole study population, as well as additional data now available after 24 months through 60 months postoperative. Patients who were secondary surgical failures were considered failures for overall success, but for the other individual variables, Last Observation Carried Forward (LOCF) was used.

The FDA will be asking you to discuss whether the data in the PMA provide reasonable assurance that the proposed device is safe for the specified indications and intended patient population and whether any additional data or analyses are needed.

CLINICAL STUDY RESULTS

Patient Accounting

The IDE study database was initially submitted for 24 month evaluation on May 15, 2006, but was resubmitted with an updated data set in April 6, 2009 in order to respond to deficiencies from FDA. The date of database closure for analyses was February 3, 2009.

The prospective, randomized controlled clinical trial included a total of 463 patients. Of these, 239 were implanted with the investigational device and 224 received the control treatment. There were 55 patients randomized who were withdrawn from the study prior to receiving their assigned treatment (23 investigational and 32 controls). These were considered as "non-study" patients and analyzed separately. There were 2 patients who were randomized to the control group but received the investigational treatment. These patients were grouped according to the device which they received.

A total of 29 sites participated in the study, but three sites withdrew prior to the study closure. From the time after surgery through 24 months postoperative, four investigational patients and 16 control patient withdrew from the study. After the 24-month time point, an additional 23 investigational patients and 17 control patients withdrew from the study. Most were reported as due to patients removing their consent for participation or sites withdrawing from the study. The number of patients who had overall success outcomes at 24 months was 200 (84.7%) investigational and 182 (82.7%) controls.

Patients were evaluated using three different analysis data sets (primary, per-protocol, and missing-equals-failure). The primary dataset consisted of all patients who received study devices with statistical comparisons based on observed data and no imputations of missing data due to lost-to-follow-up. Patients who were secondary surgical failures were considered failures for overall success in regards to the primary endpoint, but for the other individual variables, Last Observation Carried Forward (LOCF) was used. The PMA is based on the final analysis of the data evaluated through the 24-month visits for the whole study population, as well as additional long-term data now available up to 60 months postoperative.

The following table presents the patient accounting data for the 463 randomized, treated patients at 24 and 60 month follow-up:

Table 3 – Patient Accounting				
	24 month		60 month	
	Investigational	Control	Investigational	Control
Number of Patients Enrolled	239	224	239	224
Theoretical Follow-up	239	224	239	224
Cumulative Deaths	3	4	5	7
Failures (Cumulative)[1]	10 (16)	15 (29)	1 (26)	3 (39)
Patients Not yet overdue	0	0	9	15
Failure(s) was not Counted in the Denominator				
Expected[2]	220	191	199	163
Evaluated of Expected	195	157	130	99
Percent Follow-up (%)	88.6	82.2	65.3	60.7

[1]Patient who had a failure defined as an additional surgery with a supplemental fixation, non-elective implant removal or revision or defined as a serious device- or device/surgical-related adverse event not leading to a second surgery.
[2]Expected = Theoretical – Cumulative Deaths – Cumulative Failures – additional Patients to be Evaluated (i.e., Patients Not Yet Overdue)

While over 80% of patients in both groups remained in the study at 24 months (88.6 % of AMPLIFY™ patients and 82.2% of Control), follow-up at 60 months was only 65.3% and 60.7% for each group, respectively. Losing a large percentage of the study patients implies that the treatment groups may differ over time.

Demographics and Preoperative Characteristics

The following tables provide summary and comparisons of demographic variables, preoperative characteristics and evaluation of clinical endpoints, and surgery and discharge information between the AMPLIFY™ and Control groups.

Table 4 – Demographic Information			
Variable	**Investigational (N=239)**	**Control (N=224)**	**p-value**
Age (yr)	53.2 ± 12.1 (20-81)	52.3 ± 13.3 (18-86)	0.408
Height (in)	67.1 ± 3.9 (59-79)	66.8 ± 4.0 (58-78)	0.380
Weight (lbs)	187.2 ± 38.9 (104-362)	188.5 ± 41.6 (99-312)	0.720
Sex	108 male (45.2%) 131 female (54.8%)	95 male (42.4%) 129 female (57.6%)	0.575
Race	218 Caucasian (91.2%) 11 Black (4.6%) 3 Asian (1.3%) 4 Hispanic (1.7%) 3 Other (1.3%)	203 Caucasian (90.6%) 14 Black (6.3%) 2 Asian (0.9%) 4 Hispanic (1.8 %) 1 Other (0.4%)	0.848
Marital status	15 single (6.3%) 176 married (73.9%) 29 divorced (12.2%)	25 single (11.2%) 155 married (69.2%) 27 divorced (12.1%)	0.457

Table 4 – Demographic Information			
	4 separated (1.7%) 14 widowed (5.9%)	3 separated (1.3%) 14 widowed (6.3%)	
Education level	< High School 24 (10.0%) High School 64 (26.8%) > High School 151 (63.2%)	< High School 29 (13.1%) High School 73 (32.9%) > High School 120 (54.1%)	0.136
Worker's comp case	Yes 27 (11.3%) No 212 (88.7%)	Yes 28 (12.5%) No 196 (87.5%)	0.774
Spinal litigation case	Yes 6 (2.5%) No 233 (97.5%)	Yes 15 (6.7%) No 209 (93.3%)	0.042
Tobacco used	Yes 63 (26.4%) No 176 (73.6%)	Yes 59 (26.3%) No 165 (73.7%)	1.000
Alcohol used	Yes 90 (37.7%) No 149 (62.3%)	Yes 78 (34.8%) No 146 (65.2%)	0.562
Preop work status	Yes 83 (34.7%) No 156 (65.3%)	Yes 92 (41.1%) No 132 (58.9%)	0.180
Previous lumbar spine surgery	Yes 73 (30.5%) No 166 (69.5%)	Yes 62 (27.7%) No 162 (72.3%)	0.540
# previous lumbar spine surgeries	1: 48 (65.8%) >1: 25 (34.2%)	1: 43 (69.4%) >1: 19 (30.6%)	0.715
Diabetes	Yes 17 (7.1%) No 222 (92.9%)	Yes 27 (12.1%) No 197 (87.9%)	0.081
Liver Disease	Yes 5 (2.1%) No 234 (97.9%)	Yes 0 (0.0%) No 224 (100.0%)	0.062
Total Waddell signs	0: 219 (91.6%) 1: 15 (6.3%) 2: 5 (2.1%)	0: 209 (93.3%) 1: 9 (4.0%) 2: 6 (2.7%)	0.508
Non-narcotic medications	Yes 154 (64.7%) No 84 (35.3%)	Yes 140 (62.5%) No 84 (37.5%)	0.630
Weak narcotic medications	Yes 116(48.5%) No 123(51.5%)	Yes 116(51.8%) No 108(48.2%)	0.516
Strong narcotic medications	Yes 38(16.0%) No 200(84.0%)	Yes 41(18.4%) No 182(81.6%)	0.537
Muscle relaxants	Yes 55(23.1%) No 183(76.9%)	Yes 55(24.7%) No 168(75.3%)	0.743

The investigational and control groups are comparable in demographic and baseline characteristics, except for spinal litigation cases for which the average number of cases is significantly greater in the Control group than in the AMPLIFY™ group. Though not statistically significant, the average number of patients with liver disease was higher in the AMPLIFY™ group than Controls.

The FDA asked the sponsor to perform sensitivity analyses where spinal litigation cases that were overall failures for the primary endpoint were considered to be successes.

This analysis resulted in a posterior probability of non-inferiority of 99.5%. Further, the posterior probability of non-inferiority of overall success in the group including the spinal litigation cases was 99.9%, while in the groups excluding spinal litigation, was 99.7%.

Table 5 – Pre-operative Evaluation of Clinical Endpoints			
Variable	**Investigational (N=239)**	**Control (N=224)**	**p-value**
Preoperative ODI	49.9 ± 13.1 (28-86)	51.6 ± 13.3 (30-94)	0.173
Preoperative SF-36 PCS	27.8 ± 6.3 (15.3-48.7)	27.4 ± 6.7 (9.1-45.2)	0.509
Preoperative SF-36 MCS	43.9 ± 13.1 (13.3-68.5)	42.9 ± 12.3 (12.9-69.3)	0.386
Back Pain Score (0-20)	15.6 ± 3.5 (0-20)	15.8 ± 3.6 (0-20)	0.568
Leg Pain Score (0-20)	14.0 ± 4.8 (0-20)	14.0 ± 5.3 (0-20)	0.942
Degenerative Disease Characteristics			
A. Instability	30 (12.6%)	24 (10.7%)	
B. Osteophyte Formation	55 (23.0%)	60 (26.8%)	
C. Decreased Disc Height	143 (59.8%)	136 (60.7%)	
D. Thickening of Ligamentous Tissue	48 (20.1%)	49 (21.9%)	
E. Disc Herniation	206 (86.2%)	201 (89.7%)	
F. Facet Joint Degeneration	98 (41.0%)	106 (47.3%)	
Number of DDD Characteristics			
1	59 (24.7%)	44 (19.6%)	
2	86 (36.0%)	84 (37.5%)	
3	50 (20.9%)	50 (22.3%)	
4	24 (10.0%)	22 (9.8%)	
5	17 (7.1%)	18 (8.0%)	
6	3 (1.3%)	6 (2.7%)	

Surgical and Discharge Information

The mean operative time and blood loss are statistically different in the AMPLIFY™ and Control groups. The hospitalization times are not statistically different for the AMPLIFY™ and Control groups. The majority of patients in both groups had procedures at L4-L5 or at L5-S1. The distribution of treatment levels is comparable between the two treatment groups.

Table 6 – Surgical and Discharge Information		
	Investigational (N=239)	**Control (N=224)**
Operative Time (hrs)	2.5 ± 0.9 (0.9-5.4) median: 2.5	2.9 ± 1.0 (1.2-8.5) median: 2.9
Blood Loss (ml)	343.1 ± 264.5 (50-1850) median: 300	448.6 ± 301.7 (50-2400) median: 375
Hospital Stay (days)	4.1 ± 2.3 (1.0-23.0) Median 4.0	4.0 ± 1.9 (1.0-19.0) Median 4.0
Treatment Levels [n %)]	L2-L3: 7 (2.9) L3-L4: 26 (10.9)	L2-L3: 3 (1.3) L3-L4: 20 (8.9)

Table 6 – Surgical and Discharge Information		
	L4-L5: 121 (50.6) L5-S1: 83 (34.7) L5-L6: 2 (0.8)	L4-L5: 122 (54.5) L5-S1: 77 (34.4) L5-L6: 2 (0.9)
External Orthosis [n (%)]	Low Profile Brace: 78 (32.6) High Profile Brace: 35 (14.6) Corset: 78 (32.6) Other: 36 (15.1) None: 12 (5.0)	Low Profile Brace: 72 (32.1) High Profile Brace: 25 (11.2) Corset: 70 (31.3) Other: 44 (19.6) None: 13 (5.8)
Patient Classified as [n (%)]	Inpatient: 239 (100.0) Outpatient: 0 (0.0)	Inpatient: 224 (100.0) Outpatient: 0 (0.0)

The FDA will be asking the Panel to discuss the adequacy of the follow-up rates in each treatment group and the impact on the determination of reasonable assurance that the proposed device is safe and effective for the specified indications and intended patient population and whether any additional data or analyses are needed. A comparison of the demographic and other patient characteristics of those remaining in the study versus those that dropped out at the 24 and 60 month time points would help to elucidate and characterize any differences in the patient population over time.

SAFETY EVALUATION

The safety of the investigational device was assessed as part of the primary study endpoint by evaluating neurological maintenance or improvement, adverse events, and secondary surgical interventions. Safety was also evaluated based on the nature and frequency of adverse events which occurred in the AMPLIFY™ group, as compared to those that occurred in the Control group. Antibody test results were also considered as part of the safety evaluation.

All Adverse Events

The adverse events, as shown in the tables below, are reported from the randomized study which included 239 AMPLIFY™ patients and 224 Control patients enrolled in the multi-center clinical study. Adverse event rates presented are based on the number of patients having at least one occurrence for a particular adverse event divided by the total number of patients in that treatment group.

A total of 209 (87.4%) of AMPLIFY™ patients had at least one adverse event within 24 months versus 197 (87.9%) Control patients (rates not statistically different). A total of 756 events were reported in the AMPLIFY™ patients and 694 events were reported in the Controls. The 24-month data analysis was used as the primary endpoint. The AMPLIFY™ group had a statistically lower rate in the category of graft site related

adverse events (as expected) and non-union adverse events, 4.2% for the investigational group versus 10.3% for the control group.

Long term safety information was also gathered for 36 and 60 month time-points. Through 60 months a total of 1115 events were reported in 222 (92.9%) of the AMPLIFY™ patients and 1017 events in 210 Controls (93.8%), an additional 359 in the AMPLIFY™ group and 323 in the controls.

The following table presents a summary of the adverse events analyzed at ≤ 24 Months as Primary Endpoint:

Table 7 – Summary of Adverse Events at 24 Months

	Investigational		Control	
	Patients (%)	**Events**	**Patients (%)**	**Events**
All AEs	209 (87.4)	756	197 (87.9)	694
Implant or implant/ procedure associated AEs	21 (8.8)	22	34 (15.2)	36
SAEs	126 (52.7)	228	125 (55.8)	234
Implant- or Implant/surgical procedure-associated SAEs[1]	15 (6.3)	15	27 (12.1)	28

[1]arthritis (0I,2C); back and/or leg pain (4I,5C); malpositioned implant (3I,2C); neurological (2I, 1C); non-union (5I,18C); trauma (1I,0C)

The following two tables present a timecourse distribution to 24 and 60 months, respectively, of all adverse events that occurred during the clinical study by treatment group for the randomized subjects:

Table 8 – Time Course of All Adverse Events to 24 Months																
Adverse Event Type	Surgery		Postoperative (1 day - <4 Weeks)		6 Weeks (≥4 Wks - <9 Weeks)		3 Months (≥9 Wks - <5 Months)		6 Months (≥5 Mos- <9 Months)		12 Months (≥9 Mos- <19 Months)		24 Months ≥19 Mos- <30 Months)		# of Patients Reporting & Total adverse events (≤24 months)	
	I	C	I	C	I	C	I	C	I	C	I	C	I	C	Invest #events / #patients (% of 239)	Control #events / #patients (% of 224)
Anatomical/ Technical Difficulty	1/1														1/1 (0.4)	(0.0)
Arthritis/Bursitis			3/3	1/1	1/1	1/1	7/7	2/2	4/4 from (5/5)	3/3	6/6	6/6	3/2	6/6	24/23 (9.6)	19/17 (7.6)
Back and/or Leg Pain			18/18	8/7	11/10	5/5	15/15	13/13	21/20	30/28 from (32/30)	37/34	31/29	37/35	23/23	139/105 (43.9)	110/89 (39.7)
Cancer							1/1		2/2	1/1	3/3	1/1	3/3		9/9 (3.8)	2/2 (0.9)
Cardiovascular	2/2		45/37	43/39		2/2	4/4	3/3	2/2	7/7	15/13	9/8	4/4	3/3	72/53 (22.2)	67/54 (24.1)
Carpal Tunnel Syndrome									2/2	1/1	4/4	3/3	3/3	2/2	9/9 (3.8)	6/6 (2.7)
Death					1/1				1/1	2/2	1/1	1/1		1/1	3/3 (1.3)	4/4 (1.8)
Dural Injury	13/13	18/18	1/1												14/14 (5.9)	18/18 (8.0)
Gastro-intestinal			18/17	16/15		2/2	4/4	3/3	5/5	2/2	9/9	11/6	7/6	9/9	43/37 (15.5)	43/33 (14.7)
Graft Site Related				4/4		3/3		5/5		3/3		2/2			0.0 (0.0)	17/17 (7.6)
Implant Loosening/ Displacement/		0/0 from (1/1)						1/1			1/1	1/1			1/1 (0.4)	2/2 (0.9)
Infection			20/19	27/24	4/3	6/6	4/4	2/2	6/4	1/1	11/11	10/10	7/6	5/5	52/39 (16.3)	51/45 (20.1)
Malpositioned Implant	1/1		3/3	1/1		1/1	1/1								5/5 (2.1)	2/2 (0.9)
Neurological			9/9	7/6	2/2	9/7	19/17	14/13	19/19	17/17	20/18	14/14	16/15	13/11	85/70 (29.3)	74/60 (26.8)
Non-Union							1/1	8/8		6/6	8/8	6/6	1/1	3/3	10/10 (4.2)	23/23 (10.3)
Other	1/1		43/35	34/25	7/7	7/4	7/6	8/8	6/5	11/10	18/14	14/11	19/17	17/15	101/70 (29.3)	91/62 (27.7)

Table 8 – Time Course of All Adverse Events to 24 Months																
				from (36/26)												
Other Pain			2/2	3/3	2/2		1/1	1/1	6/6	5/5	11/11	4/4	9/9	19/18	31/29 (12.1)	32/29 (12.9)
Respiratory			8/8	7/7		1/1	1/1	1/1		1/1	5/4	3/3	3/3		17/16 (6.7)	13/12 (5.4)
Spinal Event			1/1	1/1		2/1	3/3	3/3	5/5	2/2	5/5	12/11	4/4	2/2	18/17 (7.1)	22/19 (8.5)
Trauma			2/2	3/3	2/2	8/8	8/8	7/7	13/12	16/15	36/33	19/17	30/27	17/16	91/69 (28.9)	70/59 (26.3)
Urogenital	0/0 from (1/1)		10/10	6/6	2/2	2/2	5/5	3/3	4/4	3/3	5/5	5/4	2/2	5/5	28/27 (11.3)	24/21 (9.4)
Vertebral Fracture	3/3	3/3								0/0 from (1/1)				1/1	3/3 (1.3)	4/4 (1.8)
Any AE	21/20 (8.4)	21/20 (8.9)	183/114 (47.7)	161.90 (40.2)	32/30 (12.6)	49/35 (15.6)	81/61 (25.5	74/65 (29.0)	96/72 (30.1)	111/88 (39.3)	195/107 (44.8)	152/99 (44.2)	148/93 (38.9)	126/86 (38.4)	756/209 (87.4)	694/197 (87.9)

Table 9 – All Adverse at 60 Months				
Adverse Event Type	Total (All)			
	I		C	
	Event	Patient n (%)	Event	Patient n (%)
Patient Who Had Any Adverse Event	1115	222 (92.9)	1017	210 (93.8)
Anatomical/Technical Difficulty	1	1 (0.4)	0	0 (0.0)
Arthritis/Bursitis	37	31 (13.0)	34	27 (12.1)
Back and/or Leg Pain	219	131 (54.8)	190	124 (55.4)
Cancer	15	12 (5.0)	5	5 (2.1)
Cardiovascular	101	73 (30.5)	84	63 (28.1)
Carpal Tunnel Syndrome	9	9 (3.8)	8	8 (3.6)
Death	6	6 (2.5)	7	7 (3.1)
Dural Injury	14	14 (5.9)	20	19 (8.5)
Gastro-intestinal	75	58 (24.3)	70	51 (22.8)
Graft Site Related	0	0 (0.0)	19	19 (8.5)
Implant Loosening/ Displacement/	1	1 (0.4)	2	2 (0.9)
Infection	60	45 (18.8)	64	51 (22.8)
Malpositioned Implant	5	5 (2.1)	2	2 (0.9)
Neurological	113	85 (35.6)	98	72 (32.1)
Non-Union	11	11 (4.6)	25	25 (11.2)
Other	174	89 (37.2)	147	80 (35.7)
Other Pain	58	47 (19.7)	59	45 (20.1)
Respiratory	17	16 (6.7)	18	14 (6.3)
Spinal Event-All	30	28 (11.7)	26	22 (9.8)
Spinal Event-Cervical	16	15 (6.3)	15	14 (6.3)
Spinal Event-Lumbar	13	13 (5.4)	10	10 (4.5)
Spinal Event-Thoracic	1	1 (0.4)	1	1 (0.4)
Trauma	131	92 (38.5)	104	76 (33.9)
Urogenital	37	33 (13.8)	32	28 (12.5)
Vertebral Fracture	3	3 (1.3)	4	4 (1.8)

There are five categories of adverse events in which the AMPLIFY™ group is greater than or equal to two percentage points higher than the control group: Arthritis/Bursitis, Back and/or Leg Pain, Neurological, Trauma, and Cancer. Specifically, the AMPLIFY™ device group had a higher percentage of adverse events involving Back and/or leg pain (43.9% vs. 39.7%) and Neurological symptoms (29.3% vs. 26.8%). The correlation of high rates of Back Pain and Neurological Status with primary outcome measure is unclear. In addition, spinal events, cancers and deaths are addressed in subsequent sections due to their nature and the notably increased cancer rates seen in the AMPLIFY™ group.

There are a number of patients in both groups with adverse events that are greater than 10% as follows: Back/leg pain, Cardiovascular, Gastrointestinal, Infection, Neurological, Other, Other pain, Trauma, and Urogenital symptoms. At 24 months, there are 29.3%

of adverse events in AMPLIFY™ patients categorized as Other and 27.7% in Control patients. When these events are further recategorized as Hypersensitivity or Immunological Concerns, 7.9% of these occurred in the AMPLIFY™ patients as compared to 4.0% in the Control group.

Serious Adverse Events

Serious Adverse Events (SAEs) are defined as WHO Grade 3 or 4. There are four categories of SAEs in which the AMPLIFY™ group is greater than or equal to one percentage points higher than the Control group: back and/or leg pain (10.0% vs. 8.0%), carpal tunnel syndrome (2.1% vs 0.9%), GI (8.8% vs 6.3%), and urogenital (6.7% vs 2.7%). There are six categories in which the Control group is greater than or equal to one percentage point higher than the AMPLIFY™ group: cardiovascular (17.9% vs. 15.1%), infection (7.6% vs 5.0%), non-union (8.0% vs. 2.1%), other (9.4% vs. 5.0%), spinal event (5.4% vs. 4.2%), and trauma (10.3% vs. 9.2%).

Table 10 provides a summary of all Grade 3 and 4 adverse events. A total of 126 (52.7%) AMPLIFY™ patients and 125 (55.8%) Control patients had at least one SAE within the first 24 months of the study. There were a total of 228 SAEs in the AMPLIFY™ patient group and 234 events in the Control patient group. An additional 137 events occurred in the 103 AMPLIFY™ patients group compared to 103 events for in the Controls at 60 months.

There is a high number of serious back and/or leg pain AE's in both groups, but higher in the AMPLIFY™ group (10.0 vs. 8.0%). As many of these patients are being treated for significant back and/or leg pain associated with lumbar disc disease, this is of concern. The FDA will ask the Panel to discuss whether the number of serious back and/or leg pain AEs should be considered as "failures" and used to determine overall success. We will also ask the Panel to discuss the impact on the determination of reasonable assurance that the proposed device is safe and effective for the specified indications and intended patient population and whether any additional data or analyses are needed.

Table 10 – Time Course of Serious Adverse Events to 24 Months (Accept all changes!!!)

Complication	Surgery		Postoperative		6 Weeks		3 Months		6 Months		12 Months		24 Months		# of Patients Reporting & Total adverse events (≤ 24 months)	
	I	C	I	C	I	C	I	C	I	C	I	C	I	C	Invest #events / #patients (% of 239)	Control #events / #patients (% of 224)
Arthritis/Bursitis						1/1	2/2	1/1	1/1	3/3	3/3	3/3	1/1		7/7 (2.9)	8/7 (3.1)
Back and/or Leg Pain			5/5	1/1	1/1	2/2	2/2		4/4	7/6	8/8	7/6	5/5	5/5	25/24 (10.0)	22/18 (8.0)
Cancer							1/1		2/2	1/1	3/3	1/1	3/3		9/9 (3.8)	2/2 (0.9)
Cardiovascular	1/1		27/24	33/33			1/1	1/1	2/2	5/5	11/9	7/7	3/3		45/36 (15.1)	46/40 (17.9)
Carpal Tunnel Syndrome									1/1	1/1	2/2	1/1	2/2		5/5 (2.1)	2/2 (0.9)
Death					1/1				1/1	2/2	1/1	1/1		1/1	3/3 (1.3)	4/4 (1.8)
Dural Injury			1/1												1/1 (0.4)	
Gastrointestinal			5/5	5/5			2/2		4/4		6/6	8/4	6/5	5/5	23/21 (8.8)	18/14 (6.3)
Graft Site Related										1/1						1/1 (0.4)
Infection			7/7	7/7	2/2	3/3					3/3	4/4	1/1	3/3	13/12 (5.0)	17/17 (7.6)
Malpositioned Implant			3/3	1/1		1/1									3/3 (1.3)	2/2 (0.9)
Neurological			4/4	3/3		1/1	4/4	4/4	4/4	3/3	2/2	2/2	1/1		15/15 (6.3)	13/13 (5.8)
Non-Union failure							1/1	7/7		5/5	4/4	5/5		1/1	5/5 (2.1)	18/18 (8.0)
Other			4/4	5/4	2/2	1/1	1/1	3/3	1/1	4/4	1/1	5/5	3/3	5/5	12/12 (5.0)	23/21 (9.4)
Other Pain					1/1		1/1		2/2	1/1	1/1		1/1	5/5	6/6 (2.5)	6/6 (2.7)
Respiratory			1/1			1/1		1/1		1/1	2/1	2/2	1/1		4/3 (1.3)	5/5 (2.2)
Spinal Event			1/1			2/1	3/3	3/3	2/2	2/2	3/3	7/7	2/2		11/10 (4.2)	14/12 (5.4)
Trauma				2/2	1/1	4/4		1/1	4/4	4/4	7/7	7/6	11/11	6/6	23/22 (9.2)	24/23 (10.3)
Urogenital			3/3	1/1	2/2		4/4	1/1	3/3		4/4	4/3	1/1	3/3	17/16 (6.7)	9/6 (2.7)
Vertebral Fracture	1/1														1/1 (0.4)	
Any Serious AE	**2/2 (0.8)**	**0/0 (0.0)**	**61/51 (21.3)**	**58/49 (21.9)**	**10/10 (4.2)**	**16/15 (6.7)**	**22/20 (8.4)**	**22/22 (9.8)**	**31/27 (11.3)**	**40/34 (15.2)**	**61/42 (17.6)**	**64/48 (21.4)**	**41/33 (13.8)**	**34/29 (12.9)**	**228/126 (52.7)**	**234/125 (55.8)**

Highlighted=higher by >1 percentage point in the investigational group. See AE discussion above for further information.

Cancer Events

In the randomized subject cohort, there were 20 serious adverse events noted as cancer for this trial, to date. Notably, the IDE study for AMPLIFY™ has a specific exclusion criterion for pre-existing cancers. Therefore, all cancers associated with this clinical trial are included in the analyses. A total of 9 cancer events occurred in 9 patients (3.8%) in the AMPLIFY™ group through 24 months and a total of 2 cancer events occurred in 2 Control patients (0.9%) through 24 months. There were 13 cases of cancer in 12 AMPLIFY™ patients (5.0%) and 4 cases of cancer in 4 Control patients (1.8%) in the Updated clinical study through 60 months, with an additional 2 events in the 12 AMPLIFY™ patients and one additional event in one additional Control patient noted at the 2010 IDE annual report. This brings the total number of cancer events to date to 15 events/12 AMPLIFY™ patients and 5 events/5 Control patients.

Although the rates in the investigational group tended to be higher than those in the control, they are not statistically different at the 24 month analysis. However, statistical significance is borderline between the AMPLIFY™ and Control groups when all cancer events though the 2010 Annual Review are considered (15 events/12 AMPLIFY™ patients and 5 events/5 Control patients). The posterior probability that the cancer rate is higher in the AMPLIFY group was found to be 94% using a beta-binomial model and non-informative priors. However, this analysis did not take into account the multiple cancer events in 2 patients in the AMPLIFY group. To further characterize the data, a time-to-event analysis (time to first cancer) showed a p-value for the Wilcoxon test of approximately 0.10.

Of these events, all were classified as not related to the device. There was no clear relationship to any demographic or other parameter among the AMPLIFY™ patients with cancers reported to 24 months according to gender (5 males and 4 females); time to diagnosis (range 3-39 months); age at surgery (range 40-69); and level of surgery (6 surgeries at L4-5). Brief summaries of the types of cancer cases and subsequent additional surgeries, as well as a detailed table of the events (Table 11) are as follows:

Investigational

The cancer types at 60 months included the following: one laryngeal cancer, one lung cancer, one non-Hodgkin lymphoma, one ovarian cancer, one pancreas cancer, two prostate cancers, one stomach cancer, one thyroid cancer, one ocular cancer, one leukemia, two basal cell carcinomas, and two squamous cell carcinomas. Three additional patients had 6 cancer events after 24 months.

Control

The cancer types seen in the control group at 24 months were as follows: one non-Hodgkin's lymphoma and one case of colon cancer; after 24 months there were 3 additional cancers, with one breast. one thyroid, and one squamous cell carcinoma.

Table 11 – Summary of Cancer Events to 60 Months									
Patient	**Group**	**Sex**	**Age at Surgery**	**Surgery Level**	**Cancer Type**	**Time of Diagnosis**	**Treatment**	**Outcome**	**Comments**
[redacted]	[redacted]	M	60	L4-5	Pancreatic	13m	Chemotherapy and radiation	Death (15m)	At 13m pt was dx'd with pancreatic cancer (poorly differentiated carcinoma).
[redacted]	[redacted]	F	55	L4-5	Lung	6m		Death (7m)	Had CT abd done I day post-op, lesion noted in R lobe of liver, "diffuse radiolucency in the liver related to the spleen", and questionable abnormality of R kidney. At 6m: admitted with bilateral LE weakness and bladder/ bowel dysfcn. Found to metastatic lung CA (adenocarcinoma)
[redacted]	[redacted]	M	69	L5-S1	Prostate	12m	Observation	Permanent Disability or Condition	11m post-op developed hematuria and urinary obstruction. Bx revealed adenocarcinoma, focal , Gleason's score 4, basaloid metaplasia
[redacted]	[redacted]	F	42	L4-5	Basal cell carcinoma, left shoulder/back	3m	surgical removal of lesion	Resolved	Hospitalization was not required and no other treatment was given.
					malignant melanoma, R eye	38m	"Radiation seeds" (38m), Radiation (66m)	Permanent Disability or Condition	Developed headaches during TX for melanoma and receiving meds. The patient also received additional radiation.
[redacted]	[redacted]	F	62	L3-4	Stomach	28m	Referral to Specialist, Chemotherapy, Radiation Therapy, Total gastrectomy	Pending	28m post-op underwent EGD with biopsy. 29m post-op underwent total gastrectomy, splenectomy, esophagojejunostomy, total omentectomy, and left subclavian line placement. 43m post-op underwent follow-up EGD biopsy

Table 11 – Summary of Cancer Events to 60 Months									
Patient	**Group**	**Sex**	**Age at Surgery**	**Surgery Level**	**Cancer Type**	**Time of Diagnosis**	**Treatment**	**Outcome**	**Comments**
[redacted]	[redacted]	M	66	L4-5	Chronic Myelogenous Leukemia	39m	Bone marrow aspirate and biopsy; medications	Pending	Patient has preop history of basal and squamous cell carcinoma.
					Basal cell carcinoma	5wks	Excision	Pending	Patient has preop history of basal and squamous cell carcinoma. Excision of basal cell carcinoma with complex reconstruction (5wk and 30m)
					Squamous cell carcinoma	12.5m	Excision and electrodessication	Pending	Patient has preop history of recurrent basal and squamous cell carcinoma. Curette and electrodessication of multiple lesions from multiple sites (13m, 24m, 58m, 78m). Wide excision with complex reconstruction (49m), Removal of well differentiated squamous cell carcinoma (78m).
[redacted]	[redacted]	M	65	L4-5	Lymphoma	29m	Lumpectomy, R groin and radiation therapy	Permanent disability or condition	29m: diagnosed w/ follicular lymphoma, lumpectomy performed, treated with XRT, continues to be followed by oncology
[redacted]	[redacted]	F	58	L4-5	Thyroid	38m	Total thyroidectomy	Pending	Papillary carcinoma of thyroid
[redacted]	[redacted]	M	67	L4-5	Prostate	29m	Internal radiation therapy via permanent palladium seeding	Resolved	29m: dx'd with prostate cancer (Gleason score 6, PSA pre-treatment was 4.7). Had internal radiation. Post-tx PSA 1.0.
[redacted]	[redacted]	F	40	L5-S1	Squamous cell carcinoma insitu	21m	Chemical peels and shave biopsy	Resolved	At 21m post-op developed red patches on arms and L 4th digit. Dx'd with squamous cell carcinoma in situ and treated locally.

Table 11 – Summary of Cancer Events to 60 Months									
Patient	**Group**	**Sex**	**Age at Surgery**	**Surgery Level**	**Cancer Type**	**Time of Diagnosis**	**Treatment**	**Outcome**	**Comments**
[redacted]	[redacted]	M	53	L3-4	Laryngeal (severe squamous dysplasia)	8m	Larynoscopy with biopsy. Radiation therapy (32 treatments)	Resolved	Dx'd 8m post-op with bilateral laryngeal cancer. Received radiation therapy.
[redacted]	[redacted]	F	69	L4-5	Ovarian carcinoma	22m	R oophorectomy, Chemotherapy, then L oophorectomy and hysterectomy	Death (53m)	18m post-op developed abd. pain. MRI at 22m revealed ovarian cancer. Had R ovarian tumor removed and chemo-therapy. 37m: had TAH-BSO with chemo (4 cycles of Carboplatin/Taxon and 1 cycle of Carboplatin/ Taxotene) preceding the surgery to reduce tumor size. Additional medications (Gemcitabine, Topotecan, Doxil, and Taxotere) and Hospice.
[redacted]	[redacted]	F	45	L5-S1	Breast	38m	R breast lumpectomy, lymph node dissection.	Pending	38m: patient noted breast lump on self-exam. Had lumpectomy, chemo-therapy, and radiation.
[redacted]	[redacted]	M	54	L5-6	Squamous cell carcinoma	55m	Surgical removal of skin cancer	Resolved	
[redacted]	[redacted]	F	56	L3-4	Thyroid	48m	Total thyroidectomy and right regional node dissection	Pending	

Table 11 – Summary of Cancer Events to 60 Months									
Patient	**Group**	**Sex**	**Age at Surgery**	**Surgery Level**	**Cancer Type**	**Time of Diagnosis**	**Treatment**	**Outcome**	**Comments**
[redacted]	[redacted]	M	66	L4-5	Non-Hodgkin's lymphoma (follicular small cell cleaved lymphoma)	13m	Left inguinal lymph node dissection. (13m), ERCP (20m)	Death (23m)	13m post-op patient developed fever, chills, sweats, fatigue and dx'd with NHL. Had surgery and chemo. 20m: ERCP revealed common bile duct stricture due to adjacent lymph-adenopathy. Dx'd with follicular small cell cleaved lymphoma and found to have pulmonary and hepatic lesions (c/w the lymphoma). Died at 23m of complications of NHL.
[redacted]	[redacted]	F	72	L4-5	Colon	5m	Right hemicolectomy with ileocolostomy	Resolved	5m post-op patient was dx'd with colon cancer. Underwent right hemicolectomy and ileocolostomy and also chemotherapy.

The FDA will be asking the Panel to discuss the clinical significance of the results of these evaluations of the reported cancer events, and discuss whether any additional evaluations or analyses are necessary. Specifically, does the Panel believe that there is a clinically and/or statistically significant rate of incidence of overall cancer events in patients treated with AMPLIFY™ and/or rhBMP-2 in general, as compared to the control group and expected rates in the general population? Has the sponsor provided adequate safety data to definitively address the issue of cancer, or are additional evaluations or analyses necessary?

In addition, the FDA would like the Panel to discuss the impact on the determination of reasonable assurance that the proposed device is safe for the specified indications and intended patient population and whether any additional data or analyses are needed

Detailed Information on Specific Adverse Event Categories

Patient Deaths

There were seven deaths in the investigational group and seven in the control group through the 60 month reporting time. The difference in death rates between the two groups was not statistically significant. All deaths were classified by the investigator as not related to the device.

The following table (Table 12) provides all known information on the 14 deaths reported through the 60 month period, including updated information provided in the 2010 IDE Annual Report:

Table 12 – Deaths to 60 Months

Patient	Group	Sex/ Age at Surgery	Surgery Level	AE category	Event	Time of Event/Time of Death	Relation	Treatment	Comments
[redacted]		M/60	L4-5	Cancer	Pancreatic Cancer	13m/15m	Not Related	Chemo-therapy and radiation	At 13m pt was dx'd with pancreatic cancer (poorly differentiated carcinoma). He died at 15m.
[redacted]		F/55	L4-5	Cancer	Metastatic lung cancer	6m/7m	Not Related		Had CT abd done I day post-op, lesion noted in R lobe of liver, "diffuse radiolucency in the liver related to the spleen", and questionable abnormality of R kidney. At 6m: admitted with bilateral LE weakness and bladder/ bowel dysfcn. Found to metastatic lung CA (adenocarcinoma). She died at 7m post-operative.
[redacted]	Investig.	F/41	L4-L5	Drug Toxicity	Drug Toxicity	66m/66m	Not Related	None	Patient at 47 months began to have back and bilateral leg pain. CT myelogram showed adjacent segment degeneration above L4-5. At 66m post-op sister reports died in sleep. Medical Examiner's report indicated that the primary cause of death was drug toxicity with a combination of illicit and prescription drugs.
[redacted]		F/69	L4-5	Cardio-vascular	Cardiorespiratory arrest	51m/51m	Not Related		At 11m the patient developed angina and underwent an emergency cardiac cath. At 40m she underwent decompression L2-4, hardware removal L4-5. At 51m taken to ER, intubated, CPR. Died.
[redacted]		F/60	L4-5		Coronary artery disease	45m/45m	Not Related		It was reported that the patient had died as a result of coronary artery disease (attributed to hypertension and hyperlipidemia).
[redacted]		F/46	L3-4	Stroke	Stroke	3w/5w	Not Related		3wks post-op admitted for a stroke and had 2 more while hospitalized. Died 5wks post-op due to complications from the strokes.
[redacted]		F/69	L4-5	Cancer	Metastatic ovarian cancer	22m/53m	Not Related	Omentectomy performed followed by chemotherapy. A second surgery was performed (total abdominal hysterectomy with bilateral salpingo-oophorectomy).	Pt diagnosed at 22m months with metastatic ovarian cancer (poorly differentiated, non-small cell). Died 53m post-op due to complications from ovarian cancer.

Table 12 – Deaths to 60 Months

Patient	Group	Sex/ Age at Surgery	Surgery Level	AE category	Event	Time of Event/Time of Death	Relation	Treatment	Comments
[redacted]		M/63	L4-5	Trauma	Trauma 2° to MVA	5m/5m	Not Related		5m post-op patient died from massive injuries received in an MVA
[redacted]		M/61	L5-S1	Trauma	Fall	34m/34m	Not Related		Had accidental fall in barn and died at 34m.
[redacted]		M/66	L4-5	Cancer	Non-Hodgkin's lymphoma	13m/23	Not Related	Hernia and left inguinal lymph node dissection, chemotherapy (13m), ERCP with stent placement (20m)	13m post-op patient developed fever, chills, sweats, fatigue and dx'd with NHL. Had surgery and chemo. 20m: ERCP revealed common bile duct stricture due to adjacent lymph-adenopathy. Dx'd with follicular small cell cleaved lymphoma and found to have pulmonary and hepatic lesions (c/w the lymphoma). 23m: ERCP with stent placement was done. Died at 23m of complications of NHL.
[redacted]		F/78	L4-5	Cardio-vascular	Cardiac Arrhythmia	17m/17m	Not Related		Reported death from cardiac arrhythmia at 17m.
[redacted]		M/37	L4-5	Cardio-vascular	Mass in heart	7m/7m	Not Related		At 5m Methadone OD; at 7m patient "began feeling bad", dx'd with mass in his heart and death occurred.
[redacted]		F/47	L4-5	Complications from shoulder surgery	Complications from shoulder surgery	56m/60m	Not Related		At 56 months post-op patient reportedly died from complications of shoulder surgery.
[redacted]		M/67	L5-S1	Neurological	Amyotrophic Lateral Sclerosis (ALS)	48m/51m	Not Related	Medication for sleep and depression; referral to a specialist for ALS	At 51 months postoperatively patient reportedly died from complications of ALS

Antibody Monitoring

Antibodies to rhBMP-2 and bovine Type I collagen were assessed preoperatively and at 6 weeks, 3, 6, and 12 months post-operatively. Antibodies to human Type I collagen were assessed if the antibody response to bovine Type I collagen were positive. Patient samples were analyzed with two ELISAs designed to measure antibodies specific for rhBMP-2. The original ELISA (used to test samples from previous clinical studies with rhBMP-2) uses protein G as a reagent to detect anti-rhBMP-2 antibodies in patient samples. The newer ELISA uses an anti-human immunoglobulin antibody reagent for detection. The following outline provides results for the three types of antibody studies evaluated.

1. ***Bovine Type I Collagen Antibody Results:*** The overall incidence of elevated antibody response to bovine Type I collagen was 18.8% (85/451). In the AMPLIFY™ patients, the incidence was 16.7% (39/234); in Control patients it was 21.2% (46/217). No patients who had an elevated antibody response to bovine Type I collagen exhibited an elevated antibody response to human Type I collagen. Control patients were not exposed to the rhBMP-2/CRM during surgery, indicating some possible prior exposure to bovine collagen.

2. ***rhBMP-2 Antibody Results:*** Formation of antibodies to rhBMP-2 was assessed in 451 patients (234 investigational, 217 controls). Forty-eight patients (22 investigational, 26 control) were excluded because no samples were obtained at the postoperative time intervals.

   The results for the two ELISA tests differed. With the original ELISA test, the overall incidence of elevated antibody response to rhBMP-2 was 1.8% (8/451), all in the AMPLIFY™ group (3.4%, 8/234). The results from the newer ELISA revealed a higher overall incidence of antibody response (4.4%, 20/451). The incidence in the AMPLIFY™ patients was 6.4% (15/234), including the eight patients identified by the original protein G-based ELISA. The incidence in control patients was 2.3% (5/217).

3. ***Neutralizing Antibody Results:*** A neutralizing antibody assay was run on samples from patients with positive antibodies to rhBMP-2. There were no positive neutralizing antibodies detected.

Six of the 15 AMPLIFY™ patients who exhibited a positive antibody response with the anti-human ELISA test reported fifteen serious adverse events (one nerve root impingement, one wound infection, one osteoarthritis in the knees, one hip pain from hip replacement surgery, one cervical disc protrusion, two instances of intolerance to foods and nausea, one cardiac after MVA, one gastric volvulus, one nausea and vomiting, two cataract surgeries, one snap in back, one instance of myoclonic jerks and loss of

consciousness, and one renal failure) through 60 months. None of these events was considered to be device- or device/surgery-associated.

In the five Control patients who had positive results to rhBMP-2 based on the new anti-immunoglobulin antibody-based ELISA, two patients reported seven serious adverse events. One patient reported six serious adverse events (1 abdominal pain, 1 gallstones, 1 nausea and vomiting, 1 pulmonary embolism, 1 obesity, and 1 rectal pain with hemorrhoids), and one patient reported a gastrointestinal event. Another patient reported one serious event of gasrtro esophageal reflux disease (GERD). None of these seven serious adverse events were considered device- or device/surgery-associated. No formal statistical comparison of antibody results was performed.

All Device-Related Adverse Events

The number of patients who had adverse events classified as complications and further classified as device-related complications over the first 24 months in the AMPLIFY™ group is 21/239 (8.8%), as compared to 34/224 (15.2%) in the Control group; three possibly implant-related severe adverse events were noted in the AMPLIFY™ group between 24 and 60 months post-surgery. Most of these events related to back/leg pain, malpositioned implants, neurological events, and non-unions. No inferential statistical comparison of adverse events between investigational and control groups was performed.

The following table outlines the adverse events classified as device-related complications by category for each treatment group:

Table 13 – Time Course of Device-Related Adverse Events to 24 Months

	Surgery		Post-operative		6 Weeks		3 Months		6 Months		12 Months		# of Patients Reporting & Total adverse events (≤24 months)	
Complication	I	C	I	C	I	C	I	C	I	C	I	C	**Invest #events / #patients (% of 239)**	**Control #events / #patients (% of 224)**
Arthritis/ Bursitis	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/1	0/0	1/1	0/0	2/2 (0.9)
Back and/or leg pain	0/0	0/0	1/1	0/0	0/0	1/1	0/0	0/0	1/1	3/3	2/2	1/1	4/4 (1.7)	5/5 (2.2)
Dural Injury	0/0	1/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0 (0.0)	1/1 (0.4)
Implant Displacement /Loosening	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/1	0/0	0/0	1/1	1/1	1/1 (0.4)	2/2 (0.9)
Mal-positioned Implant	0/0	0/0	3/3	1/1	0/0	1/1	1/1	0/0	0/0	0/0	0/0	0/0	4/4 (1.7)	2/2 (0.9)
Neurological	0/0	0/0	1/1	0/0	0/0	0/0	0/0	1/1	0/0	0/0	1/1	0/0	2/2 (0.8)	1/1 (0.4)
Non-Union	0/0	0/0	0/0	0/0	0/0	0/0	1/1	8/8	0/0	5/5	8/8	6/6	10/10 (4.2)	22/22 (9.8)

Table 13 – Time Course of Device-Related Adverse Events to 24 Months														
	Surgery		Post-operative		6 Weeks		3 Months		6 Months		12 Months		# of Patients Reporting & Total adverse events (≤24 months)	
Complication	I	C	I	C	I	C	I	C	I	C	I	C	Invest #events / #patients (% of 239)	Control #events / #patients (% of 224)
Trauma	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/1	0/0	0/0	0/0	1/1 (0.4)	0/0 (0.0)
Vertebral Fracture	0/0	1/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0 (0.0)	1/1 (0.4)
Any implant associated AE	0/0	2/2	5/5	1/1	0/0	2/2	2/2	10/10	2/2	9/9	12/12	9/9	22/21 (8.8)	36/34 (15.2)

Serious Device-Related Adverse Events

The number of patients who had serious adverse events classified as complications and further classified as serious device-related complications over the first 24 months in the AMPLIFY™ group is 15/239 (6.3%), as compared to 27/224 (12.1%) in the Control group. No inferential statistical comparison of adverse events between investigational and control groups was performed.

Most of these 42 serious (WHO Grade 3 or 4), implant-associated adverse events related to back/leg pain, malpositioned implants, neurological events, and non-unions.

The following table outlines the **serious** adverse events classified as device or device/procedure-related complications by category for each treatment group to 24 months:

Table 14 – Time Course of Serious Device-Related Adverse Events to 24 Months												
	Post-operative		6 Weeks		3 Months		6 Months		12 Months		# of Patients Reporting & Total adverse events (≤24 months)	
Complication	I	C	I	C	I	C	I	C	I	C	Invest #events / #patients (% of 239)	Control #events / #patients (% of 224)
Arthritis/ Bursitis	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/1	0/0	1/1	0/0 (0.0)	2/2 (0.9)
Back and/or leg pain	1/1	0/0	0/0	1/1	0/0	0/0	1/1	3/3	2/2	1/1	4/4 (1.7)	5/5 (2.2)
Malpositioned Implant	3/3	1/1	0/0	1/1	0/0	0/0	0/0	0/0	0/0	0/0	3/3 (1.3)	2/2 (0.9)
Neurological	1/1	0/0	0/0	0/0	0/0	1/1	0/0	0/0	1/1	0/0	2/2 (0.8)	1/1 (0.4)
Non-Union	0/0	0/0	0/0	0/0	1/1	7/7	0/0	5/5	4/4	5/5	5/5 (2.1)	18/18 (8.0)
Trauma	0/0	0/0	0/0	0/0	0/0	0/0	1/1	0/0	0/0	0/0	1/1 (0.4)	0/0 (0.0)
Any implant associated AE	5/5	1/1	0/0	2/2	1/1	8/8	2/2	9/9	7/7	7/7	15/15 (6.3)	28/27 (12.1)

Secondary Surgical Interventions

Some of the adverse events reported during the study required a surgical intervention subsequent to the initial surgery. Second surgical interventions were classified by the sponsor as one of five different types of procedures:

- Revision - A procedure that adjusts or in any way modifies the original implant configuration.
- Removal - A procedure that removes one or more components of the original implant configuration without replacement with the same type of trial implant.
- Supplemental fixation - A procedure in which additional spinal devices not approved as part of the protocol are placed.
- Reoperation - Any surgical procedure at the involved level that does not remove, modify, or add any original implant components.
- "Other" surgical procedures are ones that do not fit into the previously mentioned categories and may not even involve the lumbar spine.

Within 24 months of surgery, 110 (46.0%) AMPLIFY™ patients and 140 (62.5%) Control patients had a secondary surgical procedure. The sponsor considered revisions, removals, or supplemental fixation procedure as second surgery "failures." The percentage of patients requiring a second surgical intervention considered to be "failures" in the first 24 months is 8.4% (20/239) in the AMPLIFY™ group and 16% (36/224) in the Control group. Overall statistical comparison between the secondary surgical groups was not conducted. The sponsor did not consider reoperations or "Others" as secondary surgery failures.

Three patients (3/23) in the AMPLIFY™ group and 6 patients (6/42) in the Control group who had elective removals were not classified as "failures". There were 28 Control patients (12.5%) who required a removal as compared to 13 (5.4%) in the AMPLIFY™ group. When analyzed separately, the lower probability of removals in the AMPLIFY™ group was superior to the Control group.

Through 60 months, 11 investigational patients underwent revisions and removals, 7 of which were considered to be "failures". Eight control patients underwent revisions, removals, and supplemental fixations, and six were classified as "failures".

The following table presents time course data on secondary surgical interventions for both treatment groups through 24 months:

Table 15 – Time Course of Secondary Surgeries to 24 Months												
	1 day-<1 mo		6 weeks		3 Months		6 Months		12 Months		# Patients Reporting (%) 24 mos	
	I	C	I	C	I	C	I	C	I	C	I	C
Revisions	2	0	0	2	0	0	0	0	0	1	4 (1.7)	4 (1.8)
Removals	1	0	0	1	2	0	2	2	4	17	13 (5.4)	28 (12.5)
Non-elective	1	0	0	1	2	0	1	1	4	17	10 (4.2)	22 (9.8)
Elective	0	0	0	0	0	0	1	1	0	0	3 (1.3)	6 (2.7)
Supplemental Fixations	0	0	0	1	0	1	0	2	5	5	6 (2.5)	9 (4.0)
Reoperations	10	4	2	7	1	1	0	0	0	0	12 (5.0)	11 (4.9)
Other	2	4	2	0	7	9	15	13	37	33	62 (25.9)	60 (26.8)

The sponsor did not consider reoperations as failures, although several involved a secondary surgery at the treated level with additional decompressive procedures. Clinically, this suggests that either the original procedure and device implantation did not include a complete decompression or the device did not function as intended. The FDA will be asking the Panel to discuss whether the data in the PMA provide reasonable assurance that the proposed device is safe for the specified indications and intended patient population based on how secondary surgeries were evaluated and whether any additional data or analyses that are needed.

Safety Evaluation Summary

A safety summary is provided in both narrative and table forms. Over 80% of both groups in this study had an adverse event. A total of 209 (87.4%) of investigational patients had at least one adverse event within 24 months versus 197 (87.9%) control patients (rates not statistically different). Through 24 months a total of 756 events were reported in the investigational patients and 694 events were reported in the controls. The 24-month data analysis was used as the primary endpoint. The investigational group had a statistically lower rate in the category of graft site related and non-union adverse events.

Severe or life-threatening adverse events occurred in more than 50% of both groups. The proportion of patients having serious adverse events was not statistically different for the investigational and control groups being 52.7% (228 events) and 55.8% (234 events) respectively. The proportion of patients with device-associated or device/surgical procedure-related adverse events was somewhat lower in the AMPLIFY™ group versus the Control group (8.8% and 15.2% respectively). There were 15 events (6.3%) in the AMPLIFY™ group which were both serious and device-related. In the Control group, there were 28 events (12.1%) which were both serious and implant-related. There was no formal statistical analysis conducted on adverse events. No inferential statistical comparison of adverse events between investigational

and control groups was performed. However, including patients with severe adverse events as failures did not change the overall success outcome of the study.

Long term safety information was also gathered through 60 months. Through 60 months, a total of 1,115 events were reported in 222 (92.9%) of the investigational patients and 1,017 events in 210 controls (93.8%), an additional 359 in the investigational group and 323 in the controls since the 24 month time point. In regards to serious adverse events, an additional 137 events in 103 investigational patients compared to 103 events 72 controls were noted at 60 months.

The primary safety concern is the increased numbers of cancer events in patients treated with AMPLIFY™ compared to the control group. In the randomized subject cohort, there were 20 serious adverse events noted as cancer for this trial, to date. At 24 months, a total of 9 cancer events occurred in 9 patients (3.8%) in the AMPLIFY™ group and a total of 2 cancer events occurred in 2 Control patients (0.9%). As of June 17, 2010, at 60 months, there were a total of 15 cancer events in 12 AMPLIFY™ patients (5.0%) and 5 cancer events in 5 Control patients (2.1%).

EFFECTIVENESS EVALUATION

Primary Endpoint (Overall Clinical Success)

Primary endpoints for the clinical investigation are composite variables for clinical success in fusion, ODI, and neurological status. Overall success is considered if all of the following conditions are met: 1. Safety, based on: neurological maintenance or improvement, no additional surgical procedure classified as a "failure"; and no serious "device associated or device/surgical procedure associated" adverse events; and 2. Effectiveness based on: fusion and pain/disability (Oswestry) improvement.

Study success is based on the AMPLIFY™ 24-month overall clinical success rate being statistically non-inferior to the Control group rate. The investigational group was found to be statistically non-inferior but not statistically superior to the control group. Bayesian statistical analyses yielded a posterior probability of non-inferiority at 24 months of 99.9%. The posterior probability of superiority was found to be 83.9%. Significantly more patients in the Control group presented as "unresolved spinal litigation case" ($p = 0.042$), possibly indicating a greater incidence of disease severity. Sensitivity analyses, where all spinal litigation cases considered as overall failures for the primary endpoint were considered to be successes, showed a posterior probability of non-inferiority of 99.5%.

The following table describes the success rates and the in each treatment group for overall clinical success as well as the components of overall clinical success including

patients with spinal litigation. Study success is evaluated based on data from the 24-month follow-up evaluation.

Table 16 – Overall Success and Individual Components at 24 Months (Analysis Window)						
	Number		Success rate # (%)		Posterior Probability	
	Invest	Control	Invest	Control	Non-inferiority (%)	Superiority (%)
Primary Endpoints						
Fusion	194	169	186 (95.9)	151 (89.3)	~100.0	99.2
Oswestry	208	183	152 (73.1)	133 (72.7)	99.0	53.7
Neurologic Status	207	183	180 (87.0)	154 (84.2)	~100.0	78.5
Overall Success	200	182	121 (60.5)	101 (55.5)	99.9	83.9
2nd surgery failure	16	28				
Associated SAE	15	24				
Secondary Endpoints						
Back Pain	207	183	192 (92.8)	174 (95.1)	99.9	17.8
Leg Pain	207	182	180 (87.0)	154 (84.6)	~100.0	74.6
SF-36 PCS	204	183	171 (83.8)	150 (82.0)	99.9	68.7
SF-36 MCS	204	183	145 (71.1)	120 (65.6)	~100.0	87.7

At 24 months following surgery, for all patients for whom any data are available, the overall success for the AMPLIFY™ group is 60.5% as compared to 55.5% overall success rate for the Control group, at which timepoint non-inferiority was achieved.

At 60 months, the overall success rate dropped to 43.9% in the AMPLIFY™ group and 35.1% in the Control group; however, a formal statistical analysis was not performed on 60 month data.

The following table presents an account of available data for determination of Overall Success analysis.

Table 17 – Data Accounting for Overall Success Analysis										
	6 months		12 months		24 months		36 months		60 months	
	Inv	Control	Inv	Control	Inv	Control	Inv	Control	Inv	Control
Number of Patients who had Overall Success Outcomes	204	189	214	197	200	182	150	140	123	111
Percent of Patients who had Overall Success Outcomes (%)	85.7	84.8	90.3	88.7	84.7	82.7	63.6	63.9	54.7	55.0

At 24 months following surgery, 84.7% of AMPLIFY™ patients and 82.7% of Control patients had complete data available for analysis of overall success . At 60 months, 123 (54.7%) AMPLIFY™ patients and 111 (55.0%) Control patients were evaluated for overall success outcomes. The overall success rate dropped to 43.9% in the AMPLIFY™ group and 35.1% in Controls. Formal statistical analysis was not performed on 60 month data.

Per-Protocol Results

Patients excluded from the per-protocol analysis had major protocol deviations (i.e., did not meet the inclusion/exclusion criteria, received the wrong study treatment, or had other major protocol deviations that could potentially affect clinical outcomes). The sponsor states that every statistical comparison for the per-protocol dataset yielded a posterior probability of non-inferiority of at least 98.5%, and the investigational group demonstrated statistical superiority for fusion (probability of superiority is 99.9%).

Missing Equals Failure Results

The sponsor performed a "missing-equals-failure" analysis (see Table 18), in which deaths, patients lost-to-follow-up, and missing observations due to various reasons were considered to be "failures". In this analysis, the clinical outcome rates in the "missing-equals-failure" analysis were lower than those observed in the clinical data. Although at 24 months, the investigational group's "missing-equals-failure" overall success rate is higher than that for the control group (50.6% vs. 45.1%), this is a surprisingly low success rate. No statistical comparisons were performed using the "missing-equals-failure" dataset.

The table below contains a comparison of the Primary Analysis, Per-Protocol Analysis and Missing Equals Failure analysis:

Table 18 – Comparison of Primary, Per-Protocol, and Missing=Failure Analyses of Overall Success

(24 month data)	Number		Success Rate # (%)		Posterior Probability	
	Invest	Control	Invest	Control	Non-inferiority (%)	Superiority (%)
Implanted	239	224				
Expected	236	220				
Primary Analysis						
Fusion	194	169	186 (95.9)	151 (89.3)	~100.0	99.2
Oswestry	208	183	152 (73.1)	133 (72.7)	99.0	53.7
Neurological	207	183	180 (87.0)	154(84.2)	~100.0	78.5
Overall Success	200	182	121(60.5)	101 (55.5)	99.9	83.9
2nd Surgery Failure	16	28				
Serious Associated AE	15	24				
Back Pain	207	183	192 (92.8)	174 (95.1)	99.9	17.8
Leg Pain	207	182	180 (87.0)	154 (84.6)	~100.0	74.6

Table 18 – Comparison of Primary, Per-Protocol, and Missing=Failure Analyses of Overall Success						
(24 month data)	**Number**		**Success Rate # (%)**		**Posterior Probability**	
	Invest	**Control**	**Invest**	**Control**	**Non-inferiority (%)**	**Superiority (%)**
SF-36 PCS	204	183	171 (83.8)	150 (82.0)	99.9	68.7
SF-36 MCS	204	183	145 (71.1)	120 (65.6)	~100.0	87.7
Per-Protocol Analysis						
Fusion	172	149	167 (97.1)	132 (88.6)	~100.0	99.9
Oswestry	184	160	135 (73.4)	117 (73.1)	98.5	52.2
Neurological	183	160	159 (86.9)	136 (85.0)	99.9	69.4
Overall Success	177	158	108 (61.0)	89 (56.3)	99.7	80.7
2nd Surgery Failure	12	22				
Serious Associated AE	12	18				
Missing=Failure Analysis						
Fusion			186 (77.8)	151 (67.4)		
Oswestry			152 (63.6)	133 (59.4)		
Neurological			180 (75.3)	154 (68.8)		
Overall Success			121 (50.6)	101 (45.1)		

Additional analyses were done to evaluate the primary endpoint, classifying certain secondary surgeries as failures (reoperations and elective removals), to evaluate overall success rates as related to certain spinal injections received for pain palliation, as well as to evaluate the relationship of undetermined serious adverse events on study success. No statistically significant differences were found when each of these variables was analyzed separately for overall success. While including reoperations, elective removals, serious AE, and spinal injections as failures separately did not change the overall conclusion of the primary endpoint, it is unclear whether the overall success rate would differ if all additional failures were included in the calculations of overall success. The FDA has requested that all elements considered as failures (reoperations, elective removals, serious AE, and spinal injections) be analyzed simultaneously in order to provide a complete analysis of the changes to the overall success primary endpoint.

The FDA will be asking the Panel to discuss the success rates (in both the AMPLIFY™ and Control groups) in the context of the following considered simultaneously as failures for the primary endpoint: 1) reoperations; 2) elective removals; 3) relationship-undetermined adverse events; and 5) spinal pain interventions such as injections. In particular, we will be asking the Panel to discuss the analyses of all of the above categories as failures simultaneously rather than each of the categories individually. The overall success rate in the missing=failure dataset was less than 50 % in the control group and barely over 50% in the investigational group. This low success rate in both groups raises concerns about the overall effectiveness of this type of fusion surgery either with or without the proposed device.

Selected Components of Overall Success

Oswestry Disability Index (ODI)

ODI success is a function of the preoperative ODI score. Postoperative improvement in ODI is classified as a clinical success when a 15 point decrease is achieved as compared with the preoperative score. At all postoperative time periods for both treatment groups, the mean ODI scores improved when compared to the preoperative scores ($p<0.001$). The mean improvement in ODI scores for the AMPLIFY™ group at 24 months postoperative is 27.0 points, which is slightly greater than the mean improvement score of 25.2 for the Control group, both of which are greater than the 15 point improvement which is considered clinically significant. The mean ODI scores for the AMPLIFY™ and Control patients at the different study periods up to 24 months are provided in Table 19.

Table 19 – Time Course of ODI Scores and Success Rates to 24 Months

		Preoperative	6 weeks	3 months	6 months	12 months	24 months
Mean ODI scores	**Investigational**	49.9	37.1	27.8	24.2	23.2	22.9
	Control	51.6	37.5	30.2	27.0	26.0	26.4
ODI success rates #success/ #total (%)	**Investigational**	--	107 (46.3)	140 (61.1)	170 (75.2)	159 (71.3)	152 (73.1)
	Control	--	93 (43.5)	128 (60.1)	149 (72.3)	150 (73.9)	133 (72.7)

At 24 months following surgery, the ODI success rate for the investigational group was 73.1 %, as compared to a 72.7% rate for the control group, with a posterior probability of non-inferiority of 99.0%.

Neurological Status

Neurological success was defined as maintenance or improvement in the four neurological assessments (motor function, sensory function, reflexes, and straight leg raise) with no new permanent neurological deficits as compared to baseline at 24 months.

Scores for each neurological subsection were expressed as a percent of the maximum possible score for that subsection. A normal parameter would have a score of 100, while an abnormal parameter would have a score of less than 100. Success was declared if the difference between the postoperative and preoperative scores was greater than or equal to zero.

The following table (Table 20) outlines overall neurological status success rates to 24 months for all randomized patients. Qualitatively, higher neurological status success rates generally occurred in the AMPLIFY™ group as compared to the Control group

although the differences are not statistically significant. At 24 months following surgery, the success rate of neurological status was 87.0% for the AMPLIFY™ group, as compared to 84.2% for the Control group. Bayesian analyses yielded a posterior probability of non-inferiority of essentially 100%.

Table 20 – Time Course of Neurological Status Success to 24 Months						
Neurological Success (Improve/Maintain)	**6 Months**		**12 Months**		**24 Months**	
	Success (%)	**Failure (%)**	**Success (%)**	**Failure (%)**	**Success (%)**	**Failure (%)**
Investigational	200 (87.3)	29 (12.7)	197 (87.6)	28 (12.4)	180 (87.0)	27 (13.0)
Control	182 (87.9)	25 (12.1)	180 (88.7)	23 (11.3)	154 (84.2)	29 (15.8)

The FDA would like the Panel to comment on the following topics related to study design, as well as the impact on the determination of reasonable assurance that the proposed device is safe and effective for the specified indications and intended patient population and whether any additional data or analyses are needed. On October 10, 2001 and December 19, 2001, the FDA expressed concerns regarding the sponsor's method to determine overall neurological success. The sponsor defines neurological success based on summation of each of the components of the 4 parameter scores converted to percentages. An alternative approach involves determining success/failure for each of the parameters without computing the sums and corresponding percentage. The sponsor was advised that this issue will need to be addressed by the Panel as part of their deliberation.

Fusion Status

Radiographic success in the groups were defined as meeting definition of fusion where fusion was defined as: clear evidence of bridging trabecular bone, a continuous bony connection from the superior transverse process to the inferior transverse process on both sides; no evidence of motion, by translational motion < 3 mm and angular motion < 5° between flexion and extension; and absence of cracking as evidenced by radiolucent line though the fusion mass. At all time points to 24 months for both groups, the fusion success rates for the AMPLIFY™ group were greater than corresponding Control group success rates. At 24 months the AMPLIFY™ success rates were statistically superior (posterior probability of superiority 99.2% based on Bayesian analyses).

Table 21 – Time Course of Fusion Status Success to 24 Months						
	6 months		**12 months**		**24 months**	
	Invest	Cont	Invest	Cont	Invest	Cont
Fusion Success (%)	155 (79.1)	115 (65.3)	182 (87.5)	151 (82.5)	186 (95.9)	151 (89.3)
Fusion Failure (%)	41 (20.9)	61 (34.7)	26 (12.5)	32 (17.5)	8 (4.1)	18 (10.7)

Secondary Effectiveness Endpoints

Leg Pain Numerical Rating Scale (NRS)

The leg pain score is the summation of a patient's pain intensity and duration values as measured on numerical rating scales. Postoperative improvement in leg pain is classified as a clinical success if the difference between the postoperative and preoperative scores was greater than or equal to zero. At all postoperative time periods for both treatment groups, the mean leg pain NRS scores improved when compared to the preoperative scores. A summary of leg pain scores is provided in Table 22.

Table 22 – Time Course of Leg Pain Scores and Success to 24 Months

		Preoperative	6 weeks	3 months	6 months	12 months	24 months
Mean leg pain scores	**Investigational**	14.0	6.1	5.6	5.8	6.1	6.2
	Control	14.0	5.6	5.8	5.9	6.3	6.7
Leg pain success rates #success/ #total (%)	**Investigational**	--	202/230 (87.8)	212/228 (93.0)	201/225 (89.3)	190/222 (85.6)	180/207 (87.0)
	Control	--	189/212 (89.2)	191/212 (90.1)	188/205 (91.7)	177/202 (87.6)	154/182 (84.6)

At 24 months, the mean leg pain score was 6.2 for patients in the AMPLIFY™ group and 6.7 for patients in the Control group. The mean improvement was 7.6 and 7.3 points for the two groups, respectively, both of which are greater than 0. Success rates for leg pain in the AMPLIFY™ and Control groups were 87.0% and 84.6% respectively. Bayesian analyses showed that the posterior probability of non-inferiority of the investigational device to the control at 24 months was essentially 100%.

Any improvement in leg pain is not necessarily clinically significant. To help us understand these data, the Panel should discuss the number of patients with an improved leg pain score (stratified by the amount of improvement), highlighting the number of patients with a clinically meaningful improvement, the number with an unchanged leg pain score, and the number with a leg pain score that is worse as compared to the preoperative time point. We would also like the Panel to discuss the impact on the determination of reasonable assurance that the proposed device is safe and effective for the specified indications and intended patient population and whether any additional data or analyses are needed.

Back Pain Numerical Rating Scale (NRS)

The back pain score is the summation of a patient's pain intensity and duration values as measured on numerical rating scales. Postoperative improvement in back pain is classified as a clinical success if the difference between the postoperative and preoperative scores was greater than or equal to zero. At all postoperative time periods

for both treatment groups, the mean back pain NRS scores improved when compared to the preoperative scores. A summary of back pain scores is provided in Table 23.

Table 23 – Time Course of Back Pain Scores and Success to 24 Months

		Preoperative	6 weeks	3 months	6 months	12 months	24 months
Mean back pain scores	**Investigational**	15.6	8.7	7.0	6.8	6.6	7.1
	Control	15.8	8.1	7.8	7.9	8.1	7.8
Back pain success rates #success/ #total (%)	**Investigational**	--	208/230 (90.4)	213/227 (93.8)	214/225 (95.1)	209/222 (94.1)	192/207 (92.8)
	Control	--	197/213 (92.5)	200/213 (93.9)	193/206 (93.7)	189/203 (93.1)	174/183 (95.1)

The mean score and mean improvement scores were similar for the two treatment groups. At 24 months postoperative, the mean back pain score was 7.1 for patients in the investigational group and 7.8 for patients in the control group, with 95.1% of randomized Control patients considered a NRS back pain success as compared to 92.8% of randomized AMPLIFY™ patients. Bayesian analyses showed that the posterior probability of non-inferiority of the investigational device to the control at 24 months was 99.9%.

Similar to leg pain, any improvement in back pain is not necessarily clinically significant. To help us understand this data, the Panel should discuss the number of patients with an improved back pain score (stratified by the amount of improvement), highlighting the number of patients with a clinically meaningful improvement, the number with an unchanged back pain score, and the number with a back pain score that is worse as compared to the preoperative time point. We would also like the Panel to discuss the impact on the determination of reasonable assurance that the proposed device is safe and effective for the specified indications and intended patient population and whether any additional data or analyses are needed.

Short Form 36 (SF-36)

The SF-36 is a multipurpose Quality of Life instrument, with 36 questions such that higher scores are indicative of higher functioning/better health. The Physical Component Summary (PCS) Score is a composite of the Physical Functioning, Role Functioning, Bodily Pain and General Health Scales. In addition, the Mental Health Component Summary (MCS) Score is a composite of the Vitality, Social Functioning, Role- Emotional and Mental Health Scales.

The following table outlines mean PCS and MCS scores for each treatment group at 24 months as well as mean PCS and MCS change from baseline data at 24 months:

Table 24 – Time Course of SF-36 Scores and Success to 24 Months

		Pre-operative	6 weeks	3 months	6 months	12 months	24 months
Mean SF-36 PCS scores	**Invest**	27.8	31.6	37.4	40.7	41.5	40.9
	Control	27.4	31.9	36.1	38.4	39.1	39.7
Mean SF-36 MCS scores	**Invest**	43.9	48.4	49.6	49.4	49.4	50.7
	Control	42.9	47.4	49.4	49.8	49.0	49.2
SF-36 PCS success rates #success/ #total (%)	**Invest**	--	152/225 (67.6)	187/226 (82.7)	190/221 (86.0)	187/220 (85.0)	171/204 (83.8)
	Control	--	142/212 (67.0)	173/210 (82.4)	177/206 (85.9)	174/201 (86.6)	150/183 (82.0)
SF-36 MCS success rates #success/ #total (%)	**Invest**	--	144/225 (64.0)	159/226 (70.4)	151/221 (68.3)	152/220 (69.1)	145/204 (71.1)
	Control	--	132/212 (62.3)	155/210 (73.8)	143/206 (69.4)	138/201 (68.7)	120/183 (65.6)

Mean PCS and MCS postoperative scores through the 24 month time-point were higher than preoperative scores for both treatment groups. The mean improvements in PCS and MCS scores from preoperative to 24 months postoperative for the AMPLIFY™ group (13.1 and 6.8 points, respectively) were comparable to the values for the Control group (12.3 and 6.3, respectively), as well as the PCS success rate of 83.8% for AMPLIFY™ versus 82.0% for Controls. The MCS success rate was greater for the AMPLIFY™ group (71.1 %) than for the Control group (65.6%), Bayesian analyses established non-inferiority of the AMPLIFY™ group at 24 months for the PCS and MCS results (99.9% and essentially 100.0%, respectively).

Any improvement in SF-36 score (PCS or MCS) is not necessarily clinically significant. To help us understand this data, the Panel should discuss the number of patients with an improved PCS and MCS score (stratified by the amount of improvement), highlighting the number of patients with a clinically significant improvement, the number with an unchanged scores, and the number with scores that are worse as compared to the preoperative time point. We would also like the Panel to discuss the impact on the determination of reasonable assurance that the proposed device is safe and effective for the specified indications and intended patient population and whether any additional data or analyses are needed.

The FDA will be asking the Panel to discuss whether the clinical data in the PMA provide reasonable assurance that the proposed device is effective for the specified indication and intended patient population and what additional data or analyses are needed.

Clinical Study Discussion

The clinical data provided in this application relates to short and mid-term safety and effectiveness data for the subject device. Long-term data, although not intended as the primary outcome measure of success, suggests a less favorable profile. There remains a population that continues to have low back pain and leg pain with neurological symptoms after surgical treatment. There is a concerning number of cancers in this study and all rhBMP-2 clinical spine studies. Recombinant BMP-2 has systemic effects, not unlike any other drug, and the medical community does not have enough information that relates to its long term pharmacological effects.

Post-market surveillance studies should be conducted at the recommendation of this Panel. There are several reasons for this recommendation. The failure of this device as a permanent implant is likely to have serious adverse health consequences, both from its use as a device, with the need for technically demanding revision surgery, and from its systemic effects, most notably on antibody production. Based on the effectiveness and adverse event data even at mid term evaluation, it seems that a significant number of patients continue to experience pain and/or neurological symptoms after treatment. The ability to determine the true clinical success is confounded by the use of clinical outcome instruments, such as leg and back pain, neurological status and the SF-36, which are not based on numerical changes widely accepted in the literature as clinically meaningful differences between treatment groups, but rather a numerical and/or summation of changes intra-comparatively between preoperative and postoperative scores. Most importantly, there is a higher number of cancers in the investigational group as compared to the control group that warrants further investigation.

The FDA will be asking the Panel to discuss whether the data in the PMA provide reasonable assurance that the proposed device is safe for the specified indications and intended patient population; and specifically whether the sponsor should conduct any additional studies, including possibly a cancer registry.

ADDITIONAL CANCER DATA

The sections below pertain to non-IDE data presented as part of the sponsor's evaluation of cancer outcomes. Two data analyses were included in the sponsor's evaluation of cancer outcomes: (1) Evaluation of Cancer in Medtronic and Wyeth Clinical Studies and (2) Medicare Study of Pancreatic Cancer.

Evaluation of Cancer in Clinical Studies

For the evaluation of cancer in clinical data, the sponsor presented pooled data from 44 clinical studies to support the PMA. Some of the studies were randomized clinical trials, others were clinical studies without randomization, single-arm designs, had randomizations other than 1:1, or included post-approval studies with longer-term follow-up in only the rhBMP-2 patients.

There were 1152 patients who received rhBMP-2 compared with 1008 who did not receive rhBMP-2 in 18 clinical studies conducted by Medtronic. There were 1006 patients who received rhBMP-2 compared with 749 who did not receive rhBMP-2 in 26 clinical studies conducted by Wyeth. The average follow-up in the Medtronic studies was 3.3 years in the rhBMP-2 group and 3.2 years in the non-rhBMP-2 group. The average follow-up in the Wyeth studies was 1.4 years in the rhBMP-2 group and 1.3 years in the non-rhBMP-2 group. The Medtronic studies were for spinal uses (all for degenerative disc disease; 15 lumbar and 3 cervical) and Wyeth studies for non-spinal uses of rhBMP-2 (12 tibia, 6 maxillofacial, 4 femur, 1 radius, 1 femur and radius, 1 humerus, and 1 iliac crest).

Twenty-three Surveillance, Epidemiology, and End Results (SEER) and 4 non-SEER malignancies were reported in Medtronic's rhBMP-2 patients. Eleven SEER and 1 non-SEER malignancies were found in Medtronic's non-rhBMP-2 patients. Wyeth reported 9 SEER and 5 non-SEER malignancies in rhBMP-2 patients and 6 SEER and 4 non-SEER malignancies in non-rhBMP-2 patients. Table 25 indicates the location of malignancies reported in the clinical studies.

Table 25 – SEER* Malignancies Report from rhBMP-2 Clinical Trials (from P050036/A11)				
	Reported from Medtronic Clinical Trials (as 12/22/08)		Reported from Wyeth Clinical Trials (as of 9/3/08)	
SEER Cancer Classification	**rhBMP-2 Arms (n=1152)**	**Non-rhBMP-2 Arms (n=1008)**	**rhBMP-2 Arms (n=1006)**	**Non-rhBMP-2 Arms (n=749)**
Brain and Other Nervous System	0	0	0	1
Breast	3	2	2	3**
Cervix Uteri	0	0	0	0
Colon and Rectum	1	1	1	0
Corpus Uteri	0	1	0	0
Esophagus	0	0	0	0
Hodgkin Lymphoma	0	1	0	0
Kaposi's Sarcoma	0	0	0	0
Kidney and Renal Pelvis	1**	0	0	0
Larynx	1	0	0	0
Leukemia	1	0	0	0
Liver and Bile Duct	1	0	0	0
Lung and Bronchus	1	0	0	0
Melanoma of the Skin	3	0	0	1
Myeloma	0	0	1	0
Non-Hodgkin Lymphoma	1	1	0	0
Oral Cavity and Pharynx	0	0	0	1
Ovary	1	0	0	0
Pancreas	3**	0	0	0
Prostate	2	2	4	0
Stomach	0	0	0	0
Testis	1	0	0	0
Thyroid	3	3	0	0
Urinary Bladder	0	0	0	0
Other (Mesothelioma)^	0	0	1	0
TOTAL	23	11	9	6

* SEER = Surveillance, Epidemiology and End Result
** Denotes pre-existing or potentially pre-existing malignancy. Only one of the three cases of pancreatic cancer in the Medtronic group and one of the three breast cancer cases in the Wyeth group were considered pre-existing or potentially pre-existing.
^ While this cancer is not considered a SEER malignancy, it is considered an invasive cancer and was included in the "All Sites' category for the SIR analysis.

Four analyses of the pooled Medtronic clinical studies, the pooled Wyeth clinical studies, and the entire sample of pooled clinical study patients are included in the PMA:

1. Unadjusted comparison of malignancy cases between rhBMP-2 patients and non- rhBMP-2 patients.
2. Time-to-event analysis of malignancy between rhBMP-2 and non-rhBMP-2 groups.
3. Comparison of malignancy incidence rates between rhBMP-2 and non-rhBMP-2 groups.
4. Standardized Incidence Ratio (SIR) analysis

Overall Results

1. Medtronic trials found 2.2% of rhBMP-2 patients and 1.2% of non-rhBMP-2 patients had malignancies (p-value for Fisher's exact test: 0.096). Wyeth trials found malignancies in 1.2% of rhBMP-2 patients and 1.3% of non-rhBMP-2 patients (p-value 0.830). The pooled data across all trials indicated that 1.7% of rhBMP-2 patients developed malignancy and 1.3% of non-rhBMP-2 patients developed malignancy while enrolled in the clinical trials (p-value 0.291).

2. Log-rank and Wilcoxon tests from Kaplan-Meier analysis and p-values from Cox proportional hazards regression did not show a statistical difference between the rhBMP-2 and non-rhBMP-2 groups (see table below). Due to the nature of time-to-event analyses, the pre-existing cases were excluded.

Table 26 – Time-to-Event Analyses of Time from Treatment to Malignancy Diagnosis Between rhBMP-2/non-rhBMP-2 Patients

Source	Malignancy Type	p-value from Kaplan-Meier Analysis		Cox PHREG Analysis	
		Log-Rank Test p-value	***Wilcoxon Test p-value***	**Cox Model p-value**	**Risk Ratio (95%CI)**
Medtronic	SEER malignancies	0.204	0.234	0.297	1.481 (0.708, 3.096)
	Total malignancies	0.169	0.136	0.258	1.502 (0.742, 3.040)
Wyeth	SEER malignancies	0.788	0.702	0.636	1.307 (0.431, 3.969)
	Total malignancies	0.635	0.414	0.779	0.880 (0.360, 2.150)
Pooled	SEER malignancies	0.182	0.441	0.236	1.449 (0.785, 2.674)
	Total malignancies	0.333	0.558	0.423	1.252 (0.723, 2.168)

3. The incidence rate (per 1000 patient-years) for malignancies among rhBMP-2 patients in Medtronic trials was 6.624 (95% CI 4.286, 9.779) compared with 3.711 (95% CI 1.915, 6.482) among non-rhBMP-2 patients, yielding a rate ratio of 1.785 (95% CI 0.906, 3.676, p-value 0.096). In Wyeth trials, the rates were 8.469 (95% CI 4.371, 14.790) among rhBMP-2 and 10.110 (95% CI 4.841, 18.600) among non-rhBMP-2 patients, yielding a rate ratio of 0.838 (95% CI 0.357, 1.999, p-value 0.679). The pooled data (for all Medtronic and Wyeth studies) indicated rates of 7.128 (95% CI 5.018, 9.825) among rhBMP-2 patients and 5.210 (95% CI 3.264, 7.888) among non-rhBMP-2 patients, with a rate ratio of 1.368 (95% CI 0.810, 2.353, p-value 0.246).

4. The SIRs (adjusted for age, race, and gender) were calculated using 1997-2001 US population for Medtronic's clinical trials can be found in the table below, as provided by the sponsor. The number of observed cases was more than one and greater than the number of expected cases for melanoma of the skin, pancreatic cancer, and thyroid cancer. Pancreatic and thyroid cancer SIRs were statistically greater than expected. Since the number of thyroid cancer cases was the same in the non-rhBMP group, it was not considered to be clinically significant. For Wyeth and the pooled trials, the number of observed cases was not more than one and greater than expected for any type of cancer or the overall number of cancers. When excluding the potentially pre-existing case of pancreatic cancer, the Medtronic SIR results were no longer statistically significant (unadjusted 95% CI: 0.745, 13.088).

Table 10. Standardized Incidence Ratio (SIR) of rhBMP-2 Patients: Analysis of Medtronic Data, Including Pre-Existing Cases					
			SIR		
Malignancy SEER Category	**Number of Expected**	**Number of Observed**	***Observed SIR****	***Unadjusted** 95% CI***	***Adjusted*** 95% CI***
All Cancer Sites	23.84	23	0.965	(0.589, 1.355)	
Brain & Other Nervous System	0.323	0	0.000	(0.000, 8.537)	(0.000, 18.684)
Breast	4.748	3	0.632	(0.152, 1.582)	(0.044, 2.434)
Cervix Uteri	0.258	0	0.000	(0.000, 10.674)	(0.000, 23.359)
Colon and Rectum	2.256	1	0.443	(0.035, 1.996)	(0.000, 3.545)
Corpus Uteri	0.861	0	0.000	(0.000, 3.204)	(0.000, 7.013)
Esophagus	0.221	0	0.000	(0.000, 12.486)	(0.000, 27.326)
Hodgkin Lymphoma	0.113	0	0.000	(0.000, 24.364)	(0.000, 53.321)
Kaposi's Sarcoma	0.036	0	0.000	(0.000, 76.170)	(0.000, 166.700)
Kidney and Renal Pelvis	0.606	1	1.650	(0.131, 7.433)	(0.001, 13.198)
Larynx	0.214	1	4.681	(0.373, 21.083)	(0.003, 37.437)
Leukemia	0.488	1	2.049	(0.163, 9.228)	(0.001, 16.386)
Liver and Bile Duct	0.210	1	4.765	(0.379, 21.420)	(0.003, 38.035)
Lung and Bronchus	3.090	1	0.324	(0.026, 1.458)	(0.000, 2.588)
Melanoma of the Skin	1.101	3	2.726	(0.658, 6.824)	(0.192, 10.500)
Myeloma	0.246	0	0.000	(0.000, 11.211)	(0.000, 24.535)
Non-Hodgkin Lymphoma	0.937	1	1.068	(0.085, 4.810)	(0.001, 8.540)
Oral Cavity and Pharynx	0.575	0	0.000	(0.000, 4.793)	(0.000, 10.489)
Ovary	0.461	1	2.170	(0.173, 9.773)	(0.001, 17.354)
Pancreas	**0.463**	**3**	**6.485**	**(1.565, 16.234)**	**(0.456, 24.978)**
Prostate	3.557	2	0.562	(0.097, 1.702)	(0.016, 2.765)
Stomach	0.290	0	0.000	(0.000, 9.504)	(0.000, 20.799)
Testis	0.134	1	7.490	(0.596, 33.738)	(0.004, 59.907)
Thyroid	**0.431**	**3**	**6.961**	**(1.680, 17.426)**	**(0.490, 26.813)**
Urinary Bladder	0.905	0	0.000	(0.000, 3.047)	(0.000, 6.668)

* SIR = the number of the observed cases divided by number of the expected cases.
** The 95% confidence intervals for SIRs are not adjusted for the number of multiple comparisons examined (multiplicity).
*** The 95% confidence intervals for SIRs of individual SEER categories are adjusted for the multiplicity.

FDA Comments on Evaluation of Cancer in Clinical Studies

1. The different study designs and study methodologies may preclude combining data across multiple studies, multiple sites of the body, and multiple sponsors. Analyses regarding poolability of the data were not presented; hence, it is unclear whether pooling the data is acceptable.

2. On average, patients in with rhBMP-2 were followed only 0.1 years (about 5 weeks) longer than non-rhBMP-2 patients and twice as many (25 vs 12) rhBMP-2 patients had cancer. With little difference in follow-up time, twice as many cancers, is concerning. While the Fisher's exact p-value was not statistically significant at the 0.05 level, it is unclear whether there is sufficient power to detect a difference between these two groups if one exists.

3. On average, patients in Medtronic studies were followed 2 years longer than patients in Wyeth studies, with twice as many (34 vs 17) cancers. The differences in follow-up may introduce bias in the outcomes, especially if most of the cancers in exposed patients in Medtronic studies were later in time (and thus, more likely to be associated with the exposure) and cancers in Wyeth studies were earlier in time (or, because of time followed, always earlier than those in Medtronic studies).
4. Wyeth's rhBMP-2 group and non-rhBMP-2 group had similar malignancy risk. However, 80% more malignancies developed in the rhBMP-2 group in the Medtronic studies. Even though the difference was not statistically significant, this is a substantial increase that is unlikely to be explained solely by the increased time the Medtronic rhBMP-2 patients continued in the studies.

5. There is no information on timing of cancer outcomes. With the current available information, whether there was a consistent increase in cancers throughout the time period studied in the Medtronic trials or if there were many malignancy cases toward the end of the 3 year study periods, cannot be determined. Note, however, that a Kaplan-Meier time-to-event analysis yielded p-values for the log-rank test which were not statistically significant.

6. Incidence rates in the non-rhBMP-2 group for Medtronic were lower than in the same group in the Wyeth trials. Since the Medtronic group was followed longer, and other incidence rates are similar to the Wyeth group, it is unclear why this rate is lower. It is unclear whether patient indication or comorbidities or location of rhBMP-2 use impact cancer rates.

7. The rate ratios of the Medtronic and Wyeth trials do not appear to be comparable. It is unclear that there is sufficient power to adequately determine differences between the Medtronic and Wyeth clinical trial groups, but the rate ratios of 1.785 in the Medtronic group and 0.838 in the Wyeth groups are indicative that differences do exist and pooling data would not be appropriate.

Medicare Study of Pancreatic Cancer

The sponsor also submitted a retrospective cohort study of fee-for-service Medicare and Medicare Part B claims data assessing pancreatic cancer after use of BMP with lumbar fusion. The study included 15,460 patients with BMP and 78,194 patients without BMP identified at the time of lumbar fusion operation.

Objective
The sponsor's stated objective was to assess whether there was an increase in pancreatic cancer among patients exposed to rhBMP-2 compared to those without

exposure during lumbar spinal fusion surgery between October 2003 and December 2005.

Enrollment and follow-up

Patients were at least 67 years old, with at least 2 years continuous enrollment in Medicare and had an ICD-9-CM or CPT-4 code for lumbar fusion surgery between October 2003 and December 2005. Patients were excluded if they had a diagnosis of pancreatic cancer in the two years prior to spinal fusion surgery. Patients were "followed" from the date of surgery until one of the following: diagnosis of pancreatic cancer, death, disenrollment from Medicare, end of study period (December 2005). Thus, patients were followed for a maximum of 27 months.

Outcomes

Use of BMP during spinal fusion was captured via ICD-9-CM 84.52.

Pancreatic cancer was captured via presence of ICD-9-CM 157.xx. Multiple definitions of pancreatic cancer were used:

1. (*Primary endpoint*) Two or more claims for pancreatic cancer on different service dates in any file type beginning at the date of index surgery through the end of follow-up AND at least one procedure code associated with therapy (gastrointestinal bypass surgery, pancreatectomy, radiation therapy, and chemotherapy)
2. Two or more claims for pancreatic cancer on different service dates in any file type beginning at the date of index surgery through the end of follow-up.
3. Any diagnosis code for pancreatic cancer in any file type beginning at the date of index surgery through the end of follow-up.

Information was collected regarding confounders, including: age at time of surgery, gender, race, length of follow up, diabetes mellitus, alcohol abuse, chronic pancreatitis, gastrectomy, and cholescystectomy. Data on other known confounders (smoking, obesity, family history of pancreatic cancer) was not available and thus not collected from the Medicare data. A medical chart review provided information on smoking prevalence. This information was used for a sensitivity analysis of the primary outcome.

Statistical Plan

Odds ratios and 95% confidence intervals (CI) for pancreatic cancer comparing BMP and non-BMP patients were estimated. A SIR with 95% CI was also estimated for the risk of pancreatic cancer among the BMP and non-BMP groups. SIRs were calculated using age and gender specific incidence rates in the Surveillance, Epidemiology, and End Results (SEER) Program data between 1973 and 2002. Cox proportional hazards regression was used to determine hazard ratios (HR) and 95% CI for time to pancreatic cancer development among BMP versus non-BMP groups and was adjusted for age at

time of surgery, gender, race, diabetes mellitus, alcohol abuse, chronic pancreatitis, gastrectomy, and cholescystectomy. The HR was recalculated using information on the prevalence of smoking in patients with and without BMP exposure from the medical record review. Poisson regression was included as a secondary analysis.

Validation

Use and type of BMP was assessed in 158 charts as well as smoking status (current, former, never smoker).

Descriptive Results

Non-BMP patients contributed an average of 1.46±0.86 years to the study compared with 1.04±0.73 years for BMP patients. BMP-exposed patients tended to be slightly younger, and were more often female, black, diabetic, and with prior cholecystectomy compared to the non-BMP group. Among the 91 patients with the primary endpoint for pancreatic cancer, 8 were in the BMP group and 83 were in the non-BMP group. The median time to pancreatic cancer was 0.86 years.

Table 27 – Pancreatic Cancer Outcomes

	93,654 patients met inclusion criteria	
Number of Patients	**BMP group**	**non-BMP group**
Included in study	15,460 (16.5%)	78,194 (83.5%)
At least one ICD-9-CM for pancreatic cancer	18	164
At least two ICD-9-CM for pancreatic cancer	14	115
At least one ICD-9-CM for pancreatic cancer and at least one procedure code for cancer therapy	15	134
At least two ICD-9-CM for pancreatic cancer OR at least one ICD-9-CM for pancreatic cancer and at least one procedure code for cancer therapy	17	153
At least two ICD-9-CM for pancreatic cancer and at least one procedure code for cancer therapy	8	83

Analytic Results

The SIR among the BMP group was 0.85 (95% CI 0.26, 1.44). The SIR in the non-BMP group was 1.71 (95% CI 1.34, 2.08).

Cox proportional hazard and Poisson models presented by the sponsor were adjusted for age, gender, race, diabetes, chronic pancreatitis, alcoholism, and prior cholecystectomy.

The adjusted HR for the primary endpoint of pancreatic cancer comparing BMP to non-BMP exposure was 0.70 (95% CI 0.34, 1.45). The rate of pancreatic cancer was higher

among males, blacks, patients with diabetes, and patients 70-79 years old. HR associated with other outcome definitions are listed below.

Table 8-2. Factors Associated with Pancreatic Cancer in Multivariate Cox Regression Analysis among the Cohort that Included the Original Study Cohort Plus Patients with at Least One Year but Less Than Two Years Data Prior to Surgery

	Hazard Ratio (95% CI)		
Variable	Case Definition 1* (n=225)	Case Definition 2* (n = 143)	Case Definition 3* (n=101)
Age Group			
67-69	1	1	1
70-74	1.84 (1.22 - 2.76)	1.87 (1.19 - 2.94)	1.93 (1.09 - 3.40)
75-79	2.13 (1.41 - 3.21)	1.81 (1.12 - 2.91)	2.34 (1.32 - 4.16)
80-84	2.24 (1.38 - 3.63)	1.42 (0.76 - 2.64)	1.25 (0.56 - 2.81)
≥ 85	1.16 (0.41 - 3.28)	1.11 (0.33 - 3.65)	1.73 (0.51 - 5.90)
Gender			
Male	1	1	1
Female	0.78 (0.59 - 1.03)	0.74 (0.53 - 1.02)	0.53 (0.36 - 0.78)
Race			
White	1	1	1
Black	1.67 (0.95 - 2.94)	1.96 (1.05 - 3.63)	1.82 (0.84 - 3.94)
Other	1.62 (0.80 - 3.28)	1.15 (0.42 - 3.10)	1.20 (0.38 - 3.79)
Comorbid Conditions			
Diabetes	1.54 (1.17 - 2.02)	1.50 (1.09 - 2.07)	1.70 (1.16 - 2.49)
Chronic Pancreatitis	8.43 (4.93 - 14.4)	5.48 (2.54 - 11.8)	3.16 (0.99 - 10.1)
Alcoholism	1.41 (0.52 - 3.82)	1.45 (0.46 - 4.58)	0.64 (0.09 - 4.61)
Cholecystectomy	1.64 (0.88 - 3.04)	1.58 (0.73 - 3.41)	2.14 (0.93 - 4.92)
BMP treatment	0.92 (0.60 - 1.41)	0.96 (0.59 - 1.57)	0.74 (0.38 - 1.41)

* **Case Definition 1** = Any ICD-9-CM diagnosis code for pancreatic cancer in any file type beginning at the date of index surgery through the end of follow up.
Case Definition 2 = An ICD-9-CM diagnosis code for pancreatic cancer on more than one date of service.
Case Definition 3 = Two or more diagnosis codes for pancreatic cancer on different dates of service and at least one procedure code consistent with cancer therapy.

The adjusted rate of the primary endpoint of pancreatic cancer among BMP exposed patients was 0.49 (95% CI 0.24, 1.02) times that among non-BMP exposed patients. The rate ratio using the least restrictive definition of pancreatic cancer was 0.57 (95% CI 0.35, 0.92) comparing BMP to non-BMP patients.

Validation Results

Only 3 codes of the 6 included for each claim in the Medicare data were available in the validation set. In the chart review, 94.5% (n=52/55) of BMP that could be characterized was rhBMP-2. The positive predictive value for the code used to characterize BMP (ICD-9-CM 84.52) was 100% based on 57 reviewed claims.

Type of BMP was unknown in 42.4% of charts indicating BMP use (n=28 out of 66) during the study period. While BMP-2 use was indicated in 97.4% (n=37) of charts with identified BMP type, it was only identified in 56.1% (n=37) of reviewed charts.

Thirty of the 80 patients (38%) without a BMP claim actually received BMP according to the medical record. Thus, the sensitivity of claims to detect BMP use during the study period was 54.6% (95% 41.8-66.9%)

The prevalence of ever smoking among the BMP group was 27% and among the non-BMP group was 32%. The HR for the primary endpoint of pancreatic cancer comparing BMP to non-BMP exposure was 0.73 after adjusting for smoking prevalence as well as other captured covariates.

FDA Comments on Medicare Study of Pancreatic Cancer

1. *Enrollment and follow-up:* Given the short window of observation, it is unlikely that most pancreatic cancers potentially associated with BMP were captured. A longer follow-up time is recommended.

2. *Outcomes:* As the definition for a case becomes more restrictive, it is likely that true cases are not captured via the definition and true cases may have been identified as false negatives.

3. *Descriptive Results:* Table 8-1 below presents demographics and clinical characteristics for study subjects. There were differences between patients with consecutive enrollment in Medicare for at least two years who were included in the study and those who met all other enrollment criteria, but had only 1-2 years consecutive enrollment in Medicare (see Table 8-1, below) (excluded from the study). The differences in age, race, comorbidities, and at least one definition of the outcome may indicate that patients included in the study may not be representative of the patients that qualify to receive the device.

Table 8-1. Demographic and Clinical Characteristics of the Original Study Cohort and the Cohort Who Met Other Inclusion Criteria and Were Captured in the Data for at Least One Year but Less than Two Years Prior to Surgery

Variable	Original Study Cohort	Cohort with 1-2 Year Prior Data Before Surgery	P-value
Patients (N)	93,654	9,425	
Patient-years	130,516	14,022	
Mean age ± SD	74.7 ± 5.1	68.4 ± 3.9	
	N (%)	N (%)	
Age Group			
67-69	20,214 (21.6)	7,657 (81.2)	<0.0001
70-74	32,410 (34.6)	968 (10.3)	
75-79	25,800 (27.5)	530 (5.6)	
80-84	12,082 (12.9)	216 (2.3)	
≥ 85	3,147 (3.4)	54 ((0.6)	
Gender			
Male	32,173 (34.4)	3,190 (33.8)	0.324
Female	61,481 (65.6)	6,235 (66.2)	
Race			
White	88,104 (94.1)	8,652 (91.8)	<0.0001
Black	3,495 (3.7)	496 (5.3)	
Other	2,055 (2.2)	277 (2.9)	
Comorbid Conditions			
Diabetes	33,883 (36.2)	2,969 (31.5)	<0.0001
Chronic Pancreatitis	879 (0.9)	42 (0.4)	<0.0001
Alcoholism	1,287 (1.4)	124 (1.3)	0.6049
Cholecystectomy	2,847 (3.0)	175 (1.9)	<0.0001
BMP treatment	15,460 (16.5)	1,618 (17.2)	0.101
Pancreatic cancer			
Case definition 1	182 (0.2)	43 (0.5)	<0.0001
Case definition 2	129 (0.1)	14 (0.2)	0.77
Case definition 3	91 (0.1)	10 (0.1)	0.73

4. *Validation Results*
   - Although not noted by the sponsor, if unknown sources of BMP are included in the calculations, the sensitivity for the BMP code to detect BMP-2 use is only 52.6%.
   - In the overall Medicare study, the prevalence of BMP use was 16.5%, but it was only 5.2% in the validation study. Thus, the validation sample differed from the overall study sample, at least in regards to exposure status. With only 1/3 of the BMP use noted in the validation sample, it is unclear whether the validation sample was similar enough to the overall study sample to provide an adequate validation of the study.

POST-APPROVAL STUDY

NOTE TO PANELISTS: The FDA's inclusion of a section/discussion on a Post-Approval study (PAS) in this document should not be interpreted to mean that the FDA has made a decision on the approvability of this PMA device. The presence of post-approval study plans or commitments does not in any way alter the requirements for pre-market approval and a recommendation from the Panel on whether to approve a device or not must be based on the pre-market data. The pre-market data must reach the threshold for providing reasonable assurance of safety and effectiveness before the device can be found approvable and any post-approval study could be considered. The issues noted below are the FDA's comments regarding potential post-approval studies should the FDA find the device approvable, based on the Panel's discussions and deliberations of the pre-market data.

At the request of FDA, the sponsor submitted a post-approval study to assess the potential increased risk of cancer associated with use of spinal devices containing rhBMP-2.

FDA has worked interactively with the sponsor to develop a draft post-approval study protocol to address this concern. The sponsor has proposed a retrospective claims study of Medicare data similar to the study of pancreatic cancer outcomes presented as part of the premarket data. The proposed study includes the same 15,460 patients with BMP and 78,194 patients without BMP identified at the time of lumbar fusion operation. FDA has fundamental concerns with the proposed study. FDA has informed the sponsor of these concerns with the study proposal and continues to work interactively with the sponsor to develop an agreed on study protocol.

Objective

The sponsor's stated objective for the proposed post-approval study is to investigate whether the numerical difference of cancers observed in clinical trials utilizing rhBMP-2 is indicative of an increase in the risk of cancer among lumbar spine fusion patients exposed to rhBMP-2, as compared to patients without exposure to rhBMP-2.

Enrollment and Follow-up

Patients were identified for enrollment into the study based on having claims for lumbar fusion between October 2003 and December 2005, over age 67, and enrolled in Medicare for at least two years with no claims of cancer diagnosis or treatment. The proposed study period ends in December 2005. Patients included in the study are

"followed" from the time of surgery until diagnosis of cancer, death, disenrollment in Medicare, or December 2005.

Outcomes

The following cancers are included as part of the outcome in the proposed study: head and neck, esophagus, stomach, small bowel, colon and rectum, liver and biliary, lung, bone and soft tissue, melanoma, breast, uterus, cervix, ovary, prostate, bladder, kidney and urinary collecting system, brain, thyroid.

Four outcome measures for cancer will be included. The primary measure is listed last. A full listing of the ICD-9-CM Diagnosis Codes and CPT-4 Procedure Codes can be found in the post-approval study summary in the Panel Pack.

1. Any ICD-9-CM diagnosis code for a given cancer in any file type beginning at the date of index surgery through the end of follow up.
2. An ICD-9-CM diagnosis code for the same type of cancer on more than one date of service.
3. Two or more ICD-9 diagnosis codes for the same type of cancer on different dates of service OR at least one ICD-9 diagnosis code and one procedure code consistent with cancer therapy.
4. Two or more ICD-9 diagnosis codes for the same type of cancer on different dates of service AND at least one procedure code consistent with cancer therapy.

Statistical Plan

The sponsor proposes use Cox proportional hazards regression as the primary analysis method, adjusting for covariates including demographic factors and comorbidities such as age, race, gender, diabetes mellitus, alcoholism, and any other confounding factors identifiable in the Medicare database. The sponsor also proposes estimating standardized incidence ratios (SIR), along with 95% confidence intervals for risk of cancers in the BMP-exposed and unexposed patients compared to the SEER data.

Validation

No validation was proposed for cancer outcome codes, lumbar fusion codes, or use of BMP.

FDA Comments on Proposed Post-Approval Study

1. Enrollment and follow-up

a. Patients in the proposed Medicare study did not get the same device from this PMA, but rather a device with only 1/3 amount of the rhBMP-2.

b. The maximum follow-up in the study is 27 months (October 2003-December 2005), which is insufficient to account for the induction time expected with most cancers, even shorter than the IDE study for this device.
c. Patients under age 67 are excluded from the proposed study. However, the average age of patient receiving rhBMP-2 in all of the referenced clinical trials was 47 years (median 46 years, inter-quartile range: 38-55). Thus, due to other competing risks (e.g. death), and increased comorbidities in this population, the proposed study group may not be representative of all patients getting the device. Additionally, the higher baseline risk of cancer in an older population also decreases the power of the study to detect a difference in cancer risk with use of rhBMP-2, if an increased risk exists.

2. *Outcomes:*
   a. The following SEER malignancies are not included in the proposed study: corpus uteri, Hodgkin lymphoma, Kaposi's sarcoma, larynx, leukemia, myeloma, non-Hodgkin lymphoma, oral cavity and pharynx, pancreas, and testis. While it is unclear whether all of these cancers should be included in a proposed post-approval study, pancreatic cancer is of particular concern given the data presented to FDA.

3. *Statistical plan:*
   a. Power calculations were based on percent of expected cancer cases given the number of patients followed and the average time followed. However, the primary statistical method is longitudinal in nature (which uses rates not percents) and includes adjustment for multiple covariates, which have not been taken into account in the analysis plan.

4. *Validation:*
   a. With no validation of cohort inclusion (lumbar fusion), exposure (BMP), or outcome (cancers), it is unclear whether the claims data is sufficiently accurate to answer the proposed study question. Due to the concerns with the prior validation study, an adequate validation of the post-approval study data is vital.

The FDA would like the Panel to comment on several topics related to the proposed PAS design, including types of study designs, types of cancers to be evaluated in the PAS, and duration of the PAS. Again, the proposed PAS and the following topics related to the PAS will apply only if the device is found approvable.