



Executive Summary

Introduction

This is an Executive Summary for the Medtronic Sofamor Danek Prestige Cervical Disc (P060018). The device has been reviewed by the Orthopedic Spinal Devices Branch of the Division of General, Restorative, and Neurological Devices at the Center for Devices and Radiological Health of the Food and Drug Administration.

The Executive Summary contains an identification of the applicant and manufacturer, indications for use and contraindications, and FDA's summary review memo of the device description, preclinical, and clinical information. The memo contains the following sections:

<u>Section</u>	<u>Page Number</u>
Applicant/Manufacturer Information	2
Indications for Use/Contraindications	3
Device Description	4
Mechanical (Bench) Testing	6
Particulate Injection Study	13
Explant Analysis	16
Clinical Study	
Investigational Plan	18
Statistical Analysis Plan	23
Patient Accounting and Demographics	25
Clinical Study Results	28
Effectiveness Evaluation	29
Safety Evaluation	41

Applicant/Manufacturer Information

Applicant Name and Address:

Medtronic Sofamor Danek
1800 Pyramid Place
Memphis, TN 38132

Manufacturing Facility:

Warsaw Orthopedic Inc.
2500 Sulveus Crossing
Warsaw, IN 46582

Indications for Use:

The Prestige Cervical Disc is indicated in skeletally mature patients with cervical degenerative disc disease (DDD) at one level from C3-C7. DDD is defined as intractable radiculopathy and/or myelopathy with at least one of the following items producing symptomatic nerve root and/or spinal cord compression which is documented by patient history (e.g., pain [neck and/or arm pain], functional deficit, and/or neurological deficit), and radiographic studies (e.g., CT, MRI, x-rays, etc.): 1) herniated disc, and/or 2) osteophyte formation.

Contraindications:

The Prestige Cervical Disc should not be implanted in patients with an active infection or with an allergy to stainless steel.

Device Description:

The Prestige Cervical Disc system is a two-piece device with a metal-on-metal articulation that is inserted into the intervertebral disc space at a single level using an anterior approach. The device is manufactured from type 316 stainless steel (ASTM F-138) and consists of two metal plates which interact via a ball and trough mechanism. The superior component of the implant contains the ball portion of the mechanism, and the inferior component incorporates the trough portion. The flat portion of each component, which contacts the vertebral endplate, is aluminum oxide grit blasted for bone on-growth.

Each component is affixed to the vertebral body by two bone screws through an anterior flange. The bone screws are held in place by a lock screw mechanism. In the implanted disc, the bone screws are divergent in the cephalic/caudal direction and convergent in the medial/lateral direction.

The device assembly allows the following motions *in vitro*:

Flexion/Extension	>10°
Lateral Bending	>10°
Axial Rotation	Unconstrained
Anterior/Posterior Translation	2mm



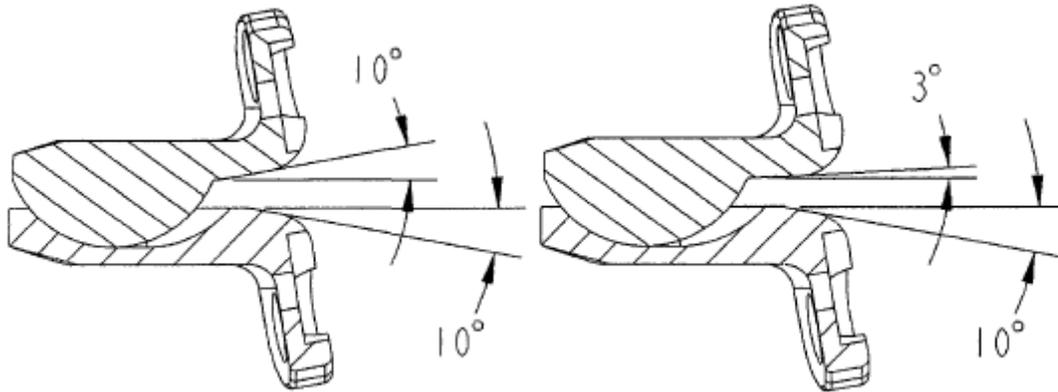
The Prestige device is available in various sizes

Discs (Height x Depth)	Self Tap Screws (Diameter x Length)
6mm x 12mm Disc	4.0mm x 13mm Bone Screw
6mm x 14mm Disc	4.0mm x 15mm Bone Screw
6mm x 16mm Disc – New Size*	4.5mm x 13mm Bone Screw
7mm x 12mm Disc	4.5mm x 15mm Bone Screw
7mm x 14mm Disc	
7mm x 16mm Disc – New Size*	Lock Screw
7mm x 18mm Disc – New Size*	
8mm x 14mm Disc	
8mm x 16mm Disc – New Size*	
8mm x 18mm Disc – New Size*	

*The sponsor has added five sizes since the IDE study (these five new sizes were not implanted in the study) based on feedback from surgeons and device usage. In addition, two sizes (8mm x 12mm and 9mm x 14mm) have been removed from the system.

Device Modifications (since the completion of patient enrollment):

The Prestige® Cervical Disc System used in the IDE Study (G010188) has a rotated cut of 10° that radiates in the anterior direction from the ball of the superior component and the trough of the inferior component. To accommodate the additional sizes requested by the surgeons (listed above), the sponsor wishes to change the cut from 10° to 3° on the superior component only. The change will be made to all implant sizes. The sponsor states that changing the cut adds mechanical strength to the implant and still allows for physiological motion. The sponsor references an article by Panjabi *et al.*¹, which states that the level with the highest flexion angle (in the implantable region: C3-C7) is C5-C6, which moves only 5.5°. The previous worst case device size with respect to ROM is the 6mm x 12mm device with a 10° cut. With a 10° cut, this device allows 13.6° of flexion. With a 3° cut, the 6mm x 12mm device allows 11.45° of flexion. The sponsor concludes that this change is acceptable because the flexion allowed by the device is still above the maximum physiologic flexion reported by Panjabi *et al.*



FDA Question for Panel:

A modification of the cut angle has been made in the superior components since the clinical study. The cut angle has been modified from 10° to 3°. This change adds material to the superior component; however, the range of motion has also been slightly decreased. The sponsor does not intend to market the 10° cut angle device, although it was the only device used in the clinical trial. Please discuss the potential impact of such a design change on the potential for impingement/function of this device and then comment on the adequacy of the clinical data collected on the original device design in addressing the safety and effectiveness of the newly proposed device design.

¹ Panjabi MM, Crisco JJ, Vasavada A, Oda T, Cholewicki J, Nibu K, Shin E. Mechanical properties of the human cervical spine as shown by three-dimensional load-displacement curves. *Spine* 2001; 26:2692-2700.

Mechanical (Bench) Testing:

The following bench tests were performed on the PRESTIGE® Cervical Disc: Static Compression, Compression Fatigue, Subluxation, Subsidence, Push-out, Pull-out and Wear Testing.

Static Compression (Test Report: TS00-059):

Completed September 26, 2000

Worst Case Design:

The sponsor determined the 6mm x 16mm device to be the worst case design for compression *fatigue* testing. However, the 8mm x 12mm device and the 8mm x 14mm device were utilized for static testing.

FDA Comment:

The 8mm x 14mm device may have been the worst case device design at the time this testing was performed. Using the sponsor's logic that the 6mm x 16mm device is worst case for dynamic compression testing makes the 6mm x 16mm device worst case for static compression testing as well. Additional testing utilizing the worst case 6mm x 16mm design may not be necessary in this case because of the results from the static testing of the 8mm x 14mm device as well as the results from the subsequent fatigue testing.

Acceptance Criteria:

The fatigue load must be greater than the compressive load on the cervical spine (74N) as reported by White and Panjabi².

Methods:

Testing was performed on both the 8mm x 12mm and the 8mm x 14mm device sizes. Three discs of each size were tested. Loading was applied at 0.1mm per second. Testing was performed with UHMWPE test blocks in order to utilize the bone screws.

Results:

Results are given in terms of a force at a given displacement into the polyethylene blocks. The 8mm x 12mm specimens had an average load of 1,343 ± 191N at 2mm of displacement and 6,279 ± 173N at 5mm of displacement. The 8mm x 14mm specimens had an average load of 1,709 ± 245 at 2mm of displacement and 5,664 ± 210 at 5mm of displacement.

FDA Comment:

This testing utilized polyethylene test blocks (so that the bone screws could be used) which were less stiff than the device material (stainless steel), thus allowing the device to subside. However, the device withstood loads that were far in excess of physiological loads.

Compression Fatigue (Test Report: TS06-084):

Completed September 12, 2005

Worst Case Design:

The 6mm x 16mm disc size, which has the shortest height and longest depth, was determined by the sponsor to be the worst case in compression fatigue.

² White A, Panjabi M. *Clinical Biomechanics of the Spine* J.B. Lippincott Company. 1990.

Acceptance Criteria:

The fatigue load must be greater than the compressive load on the cervical spine (74N) as reported by White and Panjabi³.

Methods:

Three 6mm x 16mm discs were tested under a load of 225N. Loading was performed in sinusoidal load amplitude control at 10 Hz with an R value of 10. UHMWPE test blocks were used.

Results:

The three devices each experienced run-out without failure to 10 million cycles under a 225N cyclic load.

Compression fatigue testing was also performed on the 6mm x 12mm disc, 6 x 14mm disc, 8mm x 12mm disc, 8mm x 14mm disc. All of the device sizes had 10 million cycle run-outs to at least 225N except for the 8mm x 12mm discs which were only tested to 150N.

FDA Comment:

The worst case device (6mm x 16mm disc) met the acceptance criterion.

Subsidence Testing (Test Report: TS02-140):

Completed December 11, 2001

Worst case:

The worst case device chosen was 8mm x 12mm because this device has the smallest footprint (area contacting vertebral endplates) offered.

Acceptance Criteria:

The subsidence force must be greater than the maximum *in vivo* compressive load in the cervical spine (74N) as reported by White and Panjabi⁴.

Methods:

Subsidence testing was performed five times on one 8mm x 12mm device. Grade 15 foam test blocks were used to simulated bone. Axial compressive loading was applied at 0.1mm/second until the foam blocks touched, which was a distance of ~8mm.

Results:

Specimen	Yield Strength (N)	Yield Displacement (mm)	Ultimate Strength (N)	Ultimate Displacement (mm)	Stiffness (N/mm)
Mean ± SD	550 ± 20	3.05 ± 0.20	718 ± 62	8.0	363.1 ± 37.0

FDA Comment:

The yield strength (of the foam in this case) was higher than the expected physiologic loads on the device/bone based on the acceptance criterion.

^{3,4} White A, Panjabi M. *Clinical Biomechanics of the Spine* J.B. Lippincott Company. 1990.

Subluxation Testing (Test Report TR07-042):

Completed August 8, 2006

Purpose:

Testing was done to determine the amount of force required to dislocate the upper component of the disc assembly from the lower component when the disc is in the neutral position and at extreme angles of flexion, extension, and lateral bending.

Worst Case:

No worst case device was identified for this testing because all devices share the same articulation.

Acceptance Criteria:

The subluxation force must be greater than maximum *in vivo* shear load in the cervical spine (20N) as reported by White and Panjabi⁵.

Methods:

Testing was conducted using five 7mm x 14mm discs. Polyethylene test blocks were used. A 100N compressive preload was used during all testing. Shear loads were applied to the inferior test block in displacement control at a rate of 0.1mm/sec. Each of the five discs was subject to each of the loadings below (in random order to capture possible effects from previous tests).

Angle	Direction of Shear Force
0° (neutral)	-Y (medial to lateral)
10° Lateral Bend	-Y (medial to lateral)
10° Lateral Bend	+Y (medial to lateral)
10° Flexion	-X (anterior to posterior)
10° Flexion	+X (posterior to anterior)

Results:

Angle	Direction of Shear Force	Average Peak Shear Load (N) (n=5)
0° (neutral)	-Y (medial to lateral)	111.5 ± 31.5
10° Lateral Bend	-Y (medial to lateral)	108.1 ± 9.5
10° Lateral Bend	+Y (medial to lateral)	86.8 ± 25.9
10° Flexion	-X (anterior to posterior)	104.7 ± 6.7
10° Flexion	+X (posterior to anterior)	77.1 ± 9.3

Other Subluxation Tests (Test Report: TS02-138 and Test Report: TS02-175):

The sponsor performed two other subluxation tests. However, results were far more variable potentially due to the fact that the same device was used for all tests. Therefore, at the request of FDA, the above test was performed.

⁵ White A, Panjabi M. *Clinical Biomechanics of the Spine* J.B. Lippincott Company. 1990.

FDA Comment:

The devices met the acceptance criterion.

Push-out (Test Report: TS02-150):Purpose:

Testing was done to determine the push-out load of the device in the absence of screw fixation.

Worst Case:

The 8mm x 12mm disc was used for this testing as it has the smallest footprint available and therefore the minimum surface area in contact with bone.

Acceptance Criteria:

The push-out force must be greater than maximum *in vivo* shear load in the cervical spine (20N) as reported by White and Panjabi⁶.

Methods:

Testing was completed on five 8mm x 12mm specimens. Specimens were loaded between pieces of grade 15 foam bone with 100N of preload while an axial force was applied to the posterior portion of the disc at 25mm/min until 10mm was reached. Grade 15 foam is used to mimic the physical properties of natural bone.

Results:

The average push-out load for the five samples was $129 \pm 9.6\text{N}$.

FDA Comment:

The devices met the acceptance criterion.

Pull-out (Test Report: TS02-139):

Completed December 11, 2001

Purpose:

Determine the pullout load of the Prestige disc with bone screw fixation.

Worst Case:

The 8mm x 12mm disc was used for this testing as it has the smallest footprint and therefore the minimum surface area in contact with bone.

Acceptance Criteria:

The pull-out force must be greater than maximum *in vivo* shear load in the cervical spine (20N) as reported by White and Panjabi⁷.

Methods:

Each test article consisted of one male component or one female component attached to a foam block with bone screws. Specimens were subjected to static axial pullout in accordance with ASTM F1691-96. Load was applied by a cable that loops through the screw holes of the device. Load was applied at a rate of 25mm/min. The male and female components were tested separately. The metal components were reused because they were not damaged during the test; however, the foam blocks were replaced for each run.

^{6,7} White A, Panjabi M. *Clinical Biomechanics of the Spine* J.B. Lippincott Company. 1990.

Results:

After five runs, the male components had an average pull-out strength of $200 \pm 24\text{N}$ and the female components had an average pull-out strength of $251 \pm 36\text{N}$.

This same testing was also performed on the 8mm x 14mm disc. After five runs, the male components had an average pull-out strength of $191 \pm 35\text{N}$ and the female components had an average pull-out strength of $225 \pm 50\text{N}$.

FDA Comment:

This is essentially a test of the pull out strength of the bone screws. The devices met the acceptance criterion.

Wear Testing (Test Report: TS04-135)

Completed August 13, 2004

Purpose:

Testing was done to determine the long term functionality of the Prestige device.

Worst Case Device:

Test articles consisted of an upper and lower test coupon. Because all implant sizes have identical articulating geometry, there is not a worst-case size for this test. Furthermore, this test required the use of a testing coupon in lieu of a standard device to facilitate attachment to the machines and to ensure proper measurement of the weight change of the articles. The testing coupon was a disc with the same articulating geometry and surface finish as the standard parts. The coupon does not include the bone interface geometry that is part of the standard device because the test machine does not readily allow the use of these features. However, the sponsor states that these bone interface features are irrelevant for wear testing.

Acceptance Criteria:

This testing was performed to establish the wear characteristics of this device. The wear data that were generated were used to establish the parameters for the particulate injection study in rabbits. However, the components could not show any cracks as a result of the testing.

Methods:

Two groups of three specimens each were tested in a simulator to evaluate the wear. The first group was tested in coupled lateral bending/axial rotation (LB/AR) motion followed by flexion/extension (FE). The second group was tested in the reverse order to determine the effect of motion sequence on wear. The parameters for each test are in the table below.

Motion Type	Motion/Frequency	Compressive Load	Number of Cycles
Lateral Bending/Axial Rotation	$\pm 4.7^\circ$ LB @ 2Hz coupled with $\pm 3.8^\circ$ AR at 2Hz	49N	5 million
Flexion/Extension	$\pm 9.7^\circ$ FE at 2Hz	148N	10 million

The ranges of motion (ROM) represent the total ROM of adult function spine segments measured with simulated in vivo loading due to head weight.

The simulated motions were conducted in a 25% bovine serum bath of approximately 800 ml maintained at 37°C. The test was stopped at 0.5 million cycles (Mc) at 1.0 Mc and then at a minimum of once every seven days (At 2Hz, stopping every 7 days works out to stopping about every 1.2Mc) for device cleaning, weighing and photographing. The serum was changed at each stoppage and the used serum was stored.

Results:

Volumetric wear after 15Mc

LB/AR then FE		FE then LB/AR	
Specimen	Volumetric Wear (mm ³)	Specimen	Volumetric Wear (mm ³)
SS-1	4.481	SS-4	5.152
SS-2	2.201	SS-5	2.609
SS-3	4.416	SS-6	3.804
Mean	3.699 ± 1.298	Mean	3.855 ± 1.272

The average volumetric wear rate for devices tested in 5 million cycles of LB/AR followed by 10 million cycles of FE (n=3) was 0.533 ± 0.208mm³/million cycles (for the 5 million cycles of LB/AR) and 0.067 ± 0.015mm³/million cycles (for the 10 million cycles of FE).

The average volumetric wear rate for devices tested in 10 million cycles of FE followed by 5 million cycles of LB/AR (n=3) was 0.006 ± 0.005mm³/million cycles (for the 10 million cycles of FE) and 0.733 ± 0.252mm³/million cycles (for the 5 million cycles of LB/AR).

FDA Comment:

This was a characterization test and therefore no specific acceptance criterion was identified. However, the results of the wear testing do compare favorably to certain cleared metal-on-metal hip systems. The results of this testing were used to identify the appropriate dosages for the particulate injection study in rabbits. Tested components did not show any cracks.

Additional Wear Testing (Test Report: TS02-154):

Completed May 23, 2002

This was a preliminary wear test done using two discs (FE-4 and FE-6). Similar loads were used to the above test. The flexion and extension testing was performed with a 20° of motion to 10 million cycles under a 148N load. The coupled motion testing was done with a 10.4° of lateral bending and 7.6° of axial rotation to 5 million cycles under a 49N load.

Results:

Total weight loss for specimen FE-4 was 0.00050g in flexion extension and 0.43888g in lateral bending/axial rotation. Volumetric wear was 0.063mm³ and 5.520mm³, respectively. Total weight loss for specimen FE-6 was 0.00050g in flexion/extension and 0.04998g in lateral bending/axial rotation. Volumetric wear was 0.063mm³ and 6.287mm³, respectively.

FDA Comment:

The results of this testing were similar to the results of the subsequent wear test performed by the sponsor (reported above).

FDA Question for Panel:

Please discuss the adequacy of the preclinical testing as provided by the sponsor as an adequate assessment of the long term function and durability of the Prestige device. Are any additional tests recommended?

Particulate Injection Study

The Prestige device is fabricated exclusively from ASTM F-138 stainless steel. This alloy has a nominal composition of 65.4% iron, 18% chromium, 14% nickel, and 2.6% molybdenum. Other elements are allowed at low levels. ASTM F-138-03 states that “No known surgical implant material has ever been shown to be completely free of adverse reactions in the human body, ...[and]... long term clinical experience has shown an acceptable level of biological response can be expected, if the material is used in appropriate applications.” Even so, the local effects of the particulate form of this material on periprosthetic tissues were evaluated in a rabbit model.

Prosthesis Wear Testing:

The Prestige disc was tested in custom spine simulators for a total of 20 million cycles. A cycle is defined as one complete motion in a physiologic axis. A total of 10 million flexion/extension cycles were performed. An additional 5 million right/left lateral bending cycles were performed simultaneously with 5 million cycles of axial rotation. Bovine serum from the experiment was collected for further analysis. Test samples were characterized to determine the mass loss over the course of the study. Samples that have been subjected to this wear study demonstrate wear scars that are similar in shape and location to retrieved devices, however, the extent of wear in the simulator is more pronounced than in the retrievals. The sponsor states that the wear particle size and shape are representative of particles that may be generated *in vivo* (based on wear testing); however, the particle number is a worst-case approximation.

Particle Characterization:

Three samples of bovine serum containing wear debris from two bench test specimens were centrifuged, ashed, and imaged on a scanning electron microscope at magnifications as high as 20,000X. In this analysis, a range of particle sizes was found with particle dimensions as small as 0.13 microns and as large as 1.58 microns. Five sets of particle measurements were made at 10,000X and an additional five sets of particle measurements were made at 20,000X from unique samples. Results are tabulated below.

Magnification	Mean Particle Size \pm SD (nanometers)
10,000X	554 \pm 183
10,000X	550 \pm 218
10,000X	429 \pm 146
10,000X	570 \pm 379
10,000X	595 \pm 287
20,000X	364 \pm 146
20,000X	296 \pm 112
20,000X	362 \pm 126
20,000X	399 \pm 254
20,000X	302 \pm 162

The majority of the particles were granular in shape.

Rabbit Model:

An animal model was developed to assess the local and distant response to a bolus of particles. The rabbit is the smallest common laboratory model in which this procedure can be easily conducted. The particle chemistry, shape, and size were tailored to be as close to that observed in wear tests as technically possible. The resultant metal wear debris was injected into the intervertebral space for direct contact with the spinal column. Thus, the implant site selected for this procedure mimics clinical use.

One key difference between this animal model and the clinical scenario in humans is that the dose of particles is very high and represents many years of clinical use, even for the low dose animals. Clinically, the particles would be generated gradually, whereas in this model the particles are delivered as a bolus. The particle size distribution included particles of the size range observed in previous bench testing.

Rabbit Model Methods:

This animal model was used to investigate the local and distant response to a 20-million cycle equivalent does and a 60-million cycle equivalent does of particles. The equivalent dose was determined by linearly scaling the worst-case human dose determined in the custom spine simulators to a rabbit dose based on body weight. The representative human body weight was assumed to be 75kg. This selection of human body weight is more worst-case than the body weight for an obese patient since the rabbit would receive more particles.

The selection of this dose was made based on a number of correspondences between FDA and MSD between May 2002 and November 2003. In a letter dated November 5, 2004, FDA suggested that in cases where the scaled amount of particles was less than 10,000,000 particles, that 10,000,000 particles be implanted for the high dose.

Clean particles of ASTM F138 material were obtained with a size distribution that matched the characterized spine simulator particles as closely as technically possible. In order to tailor the distribution, two separate lots of particles were blended so small and medium size particles would be appropriately represented in the distribution. Fifty percent of particles were smaller than 2.85 microns in the lot of small particles and fifty percent of particles were smaller than 9.60 microns in the lot of medium particles. The particles ranged in size from less than one micron in diameter to 44 microns in diameter. The particles used of injection were sterilized using ethylene oxide gas.

One of three doses (control, low, and high) was injected into the epidural space of each of twenty New Zealand White rabbits in a carrier of contrast media (ISOVUE M-300). Dynamic fluoroscopic video was obtained at the time of injection to confirm that the particles were delivered to the intended tissue space. The animals were euthanized at 3-months (n=9) and 6-months (n=11) time points to assess the biologic response to the particles at sites both near and distant from the site of injection.

Rabbit Model Results:

Overall animal health was good. One three-month high dose rabbit suffered a traumatic injury during a routine cage change and was euthanized 20 days following injection of the particles. The fracture was deemed to be unrelated to the test article.

There was no evidence of neurotoxicity, systemic toxicity, or local spinal effects associated with treatment with the stainless steel particles. Microscopic examination of tissues at three and six months post-epidural injection did not reveal any evidence of local or systemic lesions that were thought to be attributable to the presence of the particles. Both the low and high doses of particles were considered to be non-irritants.

Clinical Observations:

There were no observations that were considered to reflect evidence of systemic or neurotoxicity or other adverse effects directly associated with the test control article.

Necropsy and Macroscopic Observations:

The sponsor stated that there were no findings that were considered to be related to presence of the test or control material.

Clinical Pathology:

There were no changes in clinical pathology parameters in either interval for both test groups that were considered suggestive of systemic toxicity or an inflammatory response. Several parameters were noted to be statistically different from the respective control. However, the sponsor considered the occurrences spurious and due to the small group sizes for comparison rather than biological significant differences.

Histopathology:

The following conclusions were made by the sponsor:

3 months: The low- and high-dose wear debris test article did not cause any microscopic findings indicating any systemic or local toxicity three months after spinal implantation. Additional evaluation of the vertebral canal sections using an Oil-red-O stain and polarized light microscopy did not reveal any apparent wear debris. Vertebral muscle/canal and spinal cord lesions noted in one high dose rabbit were likely traumatic in nature and not test article related.

6 months: Microscopically, there were no findings indicating systemic or local toxicity by the low- and high-dose wear debris six months after spinal implantation. Additional evaluation of the vertebral canal sections using polarized light microscopy did not reveal any apparent wear debris.

The sponsor concluded that both the 3 and 6-month study intervals demonstrate that the low- and high-dose wear debris are nonirritant.

FDA Comment:

Based on the long clinical history of the material and the results of the particulate injection study, it appears that the particulate produced by the device does not cause irritation and there is no overt data to raise toxicity concerns for human use in these studies.

EXPLANT EVALUATIONS:

Three stainless steel Prestige devices were explanted and underwent histological and metallurgic analyses from third party investigators.

Histological Evaluation:

The following histological conclusions were made by the third party investigator:

“At gross, all peri-prosthetic tissue samples appeared to be small fragments of connective tissues. Staining and discoloration of the tissue samples, presumably due to metallic wear debris and corrosion, was observed in a majority of the tissue samples at gross. High-resolution radiographs showed metallic fragments as well as fine metallic debris in peri-prosthetic tissue samples from all patients. Metallic debris was found at the periphery of the tissue and in localized concentrations (foci) throughout the samples. Metallic debris was not uniformly dispersed throughout the tissues. Similar to radiographic findings, larger plate-like and round irregular metallic fragments as well as fine metallic debris were found in peri-prosthetic tissue samples from all four patients. Corrosion products were observed in many microscopic fields in tissues adjacent to the Prestige devices. Macrophages (and infrequently foreign body giant cells) were commonly seen in microscopic fields where metallic debris was found, indicative of a chronic inflammatory response. This chronic inflammatory response ranges from moderate to marked. The observed host response did not have an acute inflammatory character, or an immune component (no neutrophils, lymphocytes, plasma cells, or eosinophils). The observed chronic inflammatory response with macrophages and some foreign body giant cells represents a typical finding in peri-prosthetic tissues adjacent to metal on metal arthroplasty devices.”

Metallurgic Analysis:

The following metallurgic conclusions were made by the third party investigators:

“The concave and convex surfaces of the device had a highly polished appearance when visually examined. Even when examined with a stereomicroscope at magnification up to 60X, only a slight wear track could be observed.

Overall, the data and images collected in this study were consistent with a short term implanted total disc replacement. The retrieved implant components showed only very localized, microscopic evidence of wear that were difficult to distinguish from the as-manufactured surface. The governing wear mechanism was micro-abrasion.

The observation of biofilms, typically as “spots” on the surface, represents an incidental finding, and unrelated to the clinical performance of this implant. The biofilms were observed in the contact regions, at the periphery of contact, on the screws and in the screw holes. Biofilms are commonly observed in metal-on-metal hip replacement components, although such implants are manufactured from a different alloy.

No evidence was found of macroscopic wear or fracture of the total disc replacement components. No evidence of damage that would suggest a defect in manufacturing or processing of the implant components was identified.”

Anderson *et al.*⁸ reported on the histological results and wear assessment of the three explanted Prestige discs and concluded the following:

“Although we have described revisions of disc arthroplasty, the results are reassuring. The revision rates were low, 0.3 and 1% for the Bryan and Prestige discs, respectively, which compare favorably with those demonstrated after fusion. The indications for revision are infection and persistent radiculopathy at the level of implantation and only one failure at an adjacent level. Both prostheses could be revised by conversion to a single-level fusion without neurological complication. Wear analysis of explanted compared with simulator-tested devices showed that the wear rates were significantly lower *in vivo*, although the patterns are similar as are the characteristics of the debris. It appears that fewer than 1-M cycles in a spine simulator and perhaps only 0.1-M simulator cycles are representative of 1 year of clinical use. The histological analyses do not indicate that the short-term results were associated with significant inflammatory response.”

FDA Comment: The explant analysis appears to demonstrate that the device is behaving as expected *in vivo* and that no unexpected tissue reactions were identified.

⁸ Anderson PA, Rouleau JP, Toth JM, Riew KD. A comparison of simulator-tested and –retrieved cervical disc prostheses. *J Neurosurg* 2004; 2:202-210.

CLINICAL STUDY:

The sponsor provided the data from the prospective, controlled clinical investigations of the Prestige device. In the randomized arm, there were 276 patients who received the Prestige device and 265 patients who received the control device: an anterior plated surgical fusion utilizing bone graft and plate stabilization at 36 investigational sites. After all patients had reached the 12 month evaluation point, 137 patients in the investigational group and 148 patients in the control group who had been evaluated at 24 months, 46.4% and 46% of total patients respectively, were used in the interim analysis to evaluate the safety and effectiveness of the Prestige Cervical Disc.

No pilot study was performed prior to the initiation of the pivotal IDE clinical trial. However, there was experience outside the U.S. with the Prestige disc.

INVESTIGATIONAL PLAN FOR THE PIVOTAL STUDY

Study Design.

The study was designed as a prospective multi-center randomized trial. The assessments of safety and effectiveness of the Prestige Cervical Disc were based on 1:1 comparisons between data collected from patients with single level symptomatic cervical degenerative disc disease at one level between C3-C7, either implanted with the Prestige device or an equivalent group of patients who received an anterior plated surgical fusion utilizing bone graft and plate stabilization.

The inclusion and exclusion criteria were:

Inclusion Criteria	Exclusion Criteria
<ul style="list-style-type: none">• DDD accompanied by neck pain of discogenic origin at 1 level between C3 and C7 confirmed by history and radiographic studies. DDD was determined to be present if a herniated disc and/or osteophyte formation were noted.• At least 6 weeks unsuccessful conservative treatment or signs of progression or spinal cord/nerve root compression with continued non-operative care;• No previous surgical intervention at involved level or planned procedures at involved or adjacent levels;• ≥ 18 years of age;• Preoperative Neck Disability Index score of ≥ 30;• Preoperative neck pain score of ≥ 20 on Neck and Arm Pain Questionnaire;• Not pregnant;• Willing to sign informed consent.	<ul style="list-style-type: none">• Cervical spinal condition other than symptomatic cervical DDD requiring surgical treatment at the involved level;• Cervical instability defined by dynamic (flexion/extension) radiographs showing sagittal plane translation > 3.5 mm or sagittal plane angulation $> 20^\circ$;• > 1 cervical level requiring surgical treatment;• Fused level adjacent to the treatment level• Severe pathology of the facet joints• Prior surgical intervention at the index level;• Osteopenia or osteomalacia; Spinal metastases• Any factors associated with a diagnosis of osteoporosis (per NOF criteria)• Overt or active bacterial infection, either local or systemic; Fever (temperature $> 101^\circ\text{F}$ oral) at the time of surgery• Severe insulin dependent diabetes;• Chronic or acute renal failure or prior history of renal disease;• Allergy to the metals in the devices• Mental incompetence; Prisoner, Alcohol and/or drug abuser currently undergoing treatment;• Taking drugs interfering with bone metabolism within 2 weeks prior to the planned date of spinal surgery;• Endocrine or metabolic disorder which affects osteogenesis;• Need of postoperative medications that

	interfere with the stability of the implant; <ul style="list-style-type: none"> • Treatment with another investigational therapy within 28 days prior to or 16 weeks post planned surgery.
--	---

Post-Operative Care

The recommended post-operative care included avoidance of heavy lifting, repetitive bending, and high-impact exercise or athletic activity for 60 days postoperatively. Avoidance of prolonged NSAID use (beyond 2 weeks post-op) was also specified in the post-operative regimen. The use of electrical bone growth stimulators was prohibited during the 24-month follow-up period. Patients who smoked were also encouraged to discontinue smoking.

Evaluations

Patients were evaluated preoperatively (within 6 months of surgery), intra-operatively, and postoperatively at 6 weeks, 3, 6, 12, and 24 months, and annually thereafter until the last subject enrolled in the study had been seen for their 24 month evaluation. Complications and adverse events were evaluated over the course of the clinical trial. At each evaluation time-point, the primary and secondary clinical and radiographic outcome parameters were evaluated. Success was determined from data collected during the initial 24 months of follow-up.

Adverse Events:

Adverse events were 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. Events were defined as mild moderate severe or life threatening and association to the device or procedure was defined. Duration was classified as transient or permanent (>6 months).

For the secondary surgical intervention procedures, the protocol specifies that a supplemental fixation, device removal (except elective), or revision were classified as a treatment failure. A re-operation, elective removal or other surgical procedure was not classified as a treatment failure. All patents with secondary surgical procedures were followed for the duration of the study.

The clinical parameters assessed were functional pain/disability, neck and arm pain, general health, patient global perceived effect, doctor’s perception of results, gait, and foraminal compression test.

PRIMARY STUDY ASSESSMENTS

NDI: The Neck Disability Index (NDI)⁹ was used to measure the effects of neck pain on a patient’s ability to manage everyday life (i.e., a combined measure of pain and disability). The NDI questionnaire is based on a patient’s response to ten questions, which focus on pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping, and recreation. The responses to each question range from zero to five. A lower numeric score represents a better pain and disability status regarding that variable. A total NDI score can be determined by adding the scores of the individual questions and dividing that total by the maximum possible total score (50 if all questions are answered). This yields a percentage. Therefore, NDI scores are in a range of 0% to 100%, with a lower percentage indicating less pain and disability. The NDI questionnaire was administered preoperatively as well as at each postoperative visit. Individual patient success was based on a postoperative improvement of at least 15 points from the baseline assessment. (Pre-op Score – Post-op Score ≥ 15)

⁹ Vernon, H. and Mior, S. “The Neck Disability Index: A study of reliability and validity,” *J. Manipulative Physiol. Ther.* 1991; 14(7): 409-415.

Neurological Status of the patients participating in the clinical study was assessed preoperatively and postoperatively at every follow-up visit. The neurological status questionnaire evaluated motor and sensory function as well as reflexes. Investigators judged if the patients were “normal” for these categories and, if not, specific measurements of the abnormal findings were required. Neurological success for each of the three indicators was based on maintenance or improvement of condition postoperatively as compared to the preoperative status for each element.

FSU: The FSU height was determined from lateral neutral radiographs of the treated spinal area and was expressed in millimeters. The anterior FSU height was obtained by measuring from the anterior-most point of the endplate on the superior ventral cortical margin of the cephalic vertebral body to the anterior-most point on the inferior ventral cortical margin of the caudal vertebral body of the treated segment. The posterior FSU height was determined similarly from the posterior aspect. By comparing the magnification-corrected measurements over time, one can determine if the FSU height had changed. A notable decrease in FSU height over time is considered indicative of a decrease in disc space height. FSU height was considered to be maintained or improved, i.e., success, if either the anterior or posterior postoperative measurement was no more than 2 mm less than the 6-week postoperative measurement.

SECONDARY STUDY ASSESSMENTS

General Health Status was assessed using the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36). The SF-36 scale measures specific health concepts related to physical functioning and limitations, social functioning, and health perceptions. The questionnaire contains 36 questions that pertain to eight subscales of health status. These eight subscales are physical function, role-physical, pain index, general health perception, vitality, social function, role-emotional, and mental health. These eight SF-36 scales can be summarized into two measures pertaining to physical health and mental health. The physical health summary (PCS) is based primarily on the physical functioning, role-physical, bodily pain, and general health scales of the SF-36 survey. The mental health summary (MCS) is comprised primarily of the vitality, social functioning, role-emotional, and mental health scales. Higher scores represent higher levels of health.

Neck Pain status was assessed using numerical rating scales. Neck pain is a composite of pain intensity and duration. Success was described as: Preoperative Score – Postoperative Score ≥ 0

Arm Pain Status was assessed using numerical rating scales. Success was defined as: Preoperative Score – Postoperative Score ≥ 0

Patient Satisfaction was determined using a patient questionnaire with 3 questions. At each postoperative time point, patients were asked to respond to three statements pertaining to their satisfaction with the study treatment. These statements were as follows:

1. I am satisfied with the results of my surgery.
2. I was helped as much as I thought I would be with my surgery.
3. All things considered I would have the surgery again for the same condition.

Each statement had a series of possible responses ranging from “definitely true” to “definitely false”. Success for each question was defined as “Definitely True” or “Mostly True” responses.

Patient Global Perceived Effect was assessed with a questionnaire. Success was defined as a “Completely Recovered,” “Much Improved,” or “Slightly Improved” response.

Gait Assessment was based on Nurick's classification. Success was defined as Preoperative Score – Postoperative Score ≥ 0

Foraminal Compression Test is performed by applying a force to the top of the head while the patient laterally flexes his/her head. If the patient feels pain in the upper extremities, it is likely due to nerve root compression and is considered a “positive” result. The desirable outcome is “negative” – an absence of any sensation.

Other radiographic assessments: Other radiographic outcome parameters were based on review of plain radiographs and consisted of evaluations of motion and fusion at the treated level for the investigational and control group, respectively. Adjacent level motion was also evaluated. The radiographic review was completed by two independent radiographic reviewers. If there was disagreement regarding radiographic findings between the two reviewers, a third independent reviewer adjudicated the results.

Adjacent Level Stability was assessed by motion measurements on flexion/extension and lateral bending radiographic films at the segment above and the segment below the surgical level.

Adjacent Level Measurements: In order to determine the effect, if any, of the study treatment on adjacent levels, the stability of the cervical segments above and below the treated level was assessed. The measurements were made from flexion/extension films preoperatively and postoperatively beginning at 6 weeks.

Return to Work was measured in comparison to the patient's preoperative condition

Doctor's Perception was assessed as excellent, good, fair or poor by the treating physician.

PRIMARY STUDY ENDPOINTS/SUCCESS CRITERIA

The primary endpoint was determined at 24 months as a composite of the following parameters: pain and functional disability, neurological status, adverse events, secondary surgical interventions, and a radiographic spinal unit height determination. This was termed Overall Success:

In the approved protocol, individual subject success (i.e. overall success) was defined as attainment of all of the following:

1. An improvement of at least 15 points from the baseline Neck Disability Index score;
2. Maintenance or improvement in neurological status;
3. No serious adverse event classified as implant-associated or implant/surgical procedure-associated; and
4. No additional surgical procedure classified as “Failure.”
5. Functional spinal unit (FSU) height maintenance. FSU height was considered maintained if it did not decrease more than 2 mm after 6 weeks following surgery.

In addition radiographic success was determined but with different criteria for each study group.

Radiographic success for the investigational group was based on 1) the existence of flexion/extension angular motion in a range of $> 4^\circ$ to $\leq 20^\circ$, and 2) no evidence of bridging trabecular bone forming a continuous connection between vertebral bodies.

Radiographic success for control patients was based on the presence of fusion of the treated spinal segment. To be considered fused, there had to be radiographic evidence of bone spanning the two vertebral bodies in the treated segment. Additional criteria for fusion included

flexion/extension angular motion stability ($\leq 4^\circ$) and no radiolucent lines covering more than 50% of the implant surface.

Metal Ion Testing

Metal ion testing was not part of the original IDE study protocol. Both the sponsor and FDA agreed that the collection of this information in a limited number of patients would provide some useful information. Recent study information has suggested increased chromosomal aberrations may be associated with patients who had metal on metal hip implants as compared to those with metal on poly implants. While these studies are inconclusive, most scientists agree that further information needs to be collected. Metal ion level studies are currently being conducted in a subset of patients enrolled in the Continued Access arm of the Artificial Cervical Disc pivotal IDE clinical trial. According to the protocol, patients are required to provide blood samples preoperatively and at 3, 6, 12, and 24 months following surgery. The blood samples collected as part of this study will be analyzed at Rush University Medical Center, Chicago, Illinois, for the presence of chromium and nickel ions. Testing will be done using analytical chemistry instruments, and metal ion quantities at each postoperative evaluation will be compared to the preoperative measurement.

Patients participating in the metal ion study must meet both the inclusion and exclusion criteria for the main Continued Access study as well as specific criteria for the ion study. These criteria specific to the ion study exclude patients who have metal implants, are taking certain medications or nutritional supplements, or experience occupational exposure to metal particles. The metal ion study was approved for up to 25 patients, and all of them have been enrolled and undergone surgery. A preliminary report is included in this panel package (under "Preliminary Metal Ion Study Results" tab); however, analysis of the complete metal ion testing results will be provided to FDA in the future.

STATISTICAL ANALYSIS PLAN:

Randomization and Blinding Scheme:

Patients were randomized according to the PLAN Procedure in SAS. Treatment randomization was 1:1 on a site basis.

Both the investigator and the patient were not blinded to the randomized treatment following the assignment of treatment group by the sponsor.

Hypotheses to be tested:

This is a non-inferiority trial and a fixed non-inferiority margin of 10% was agreed upon by the sponsor and FDA. The null non-inferiority hypothesis is:

$H_0 : p_{prestige} + 0.1 \leq p_{control}$, and the alternative non-inferiority hypothesis is:

$H_\alpha : p_{prestige} + 0.1 > p_{control}$, where $p_{prestige}$ and $p_{control}$ are overall success rates for the treatment and control groups respectively.

The null non-inferiority hypothesis will be rejected if the posterior probability $P(p_{prestige} + 0.1 > p_{control} | data)$ be at least 95%. The study is deemed successful if the null non-inferiority hypothesis is rejected.

If the sponsor succeeds in claiming non-inferiority, the superiority will be tested. The null superiority hypothesis is $H_0 : p_{prestige} \leq p_{control}$, and the alternative superiority hypothesis is

$H_\alpha : p_{prestige} > p_{control}$.

The null superiority hypothesis will be rejected if the posterior probability

$P(p_{prestige} > p_{control} | data)$ be at least 95%.

All other primary sub-endpoints and secondary endpoints will be tested in a similar manner.

Bayesian interim analysis plan:

An interim analysis was planned when a total of approximately 250 patients have follow-up visits at 24 months. At that time point, all the patients are expected to have reached 12-month evaluation period. If the posterior probability $P(p_{prestige} + 0.1 > p_{control} | data)$ is at least 95%, the sponsor will develop a PMA submission regarding non-inferiority.

In fact the sponsor's calculation showed that $P(p_{prestige} + 0.1 > p_{control} | data)$ is greater than 95% when the first 250 patients had valid outcomes in overall success at 24 months. Hence this PMA application was primarily based on the pre-defined interim analysis criteria. Simulations need to be performed with this margin of 10% (non-inferiority) to show that the trial design was appropriate in the sense that the type I error rate was below 5%.

Note that non-informative or uniform priors were used in the Bayesian design for this pivotal study.

Statistical Methodology:

This is a non-inferiority trial with a margin of 10%. Bayesian methods with non-informative or uniform priors were used to obtain the posterior probabilities of non-inferiority and superiority. The Bayesian model incorporates data from both the 24-month follow-up visit and 12-month follow-up visit, including those from only the 12-month visit or only the 24-month visit. However, the main focus of the analysis is the success rates at 24 months.

A multinomial model was proposed to model the correlation between the 24-month and 12-month follow-up visits. The sponsor assumes independent uniform Dirichlet (non-informative) priors for the parameters in the multinomial model. To verify the validity of this model, the sponsor calculated the correlation between the 24-month and 12-month follow-up visits based on available data. The existence of strong correlations supported the assumption of this model.

The sponsor used three different analysis datasets (primary, per-protocol, and missing equals-failure datasets) in the statistical analysis. The primary dataset consists of all the patients who received study devices and completed surgical procedures. Primary statistical comparisons were based on this dataset; missing data due to lost to follow up were not imputed. The per-protocol dataset was a subset of primary dataset. Patients who had major protocol deviations were excluded from this dataset. This dataset was used as a secondary analysis for the primary endpoint. The missing-equals-failure dataset was also constructed for a secondary analysis for the primary endpoint. In this dataset, all missing responses were assumed to be failures.

An interim analysis was planned when a total of approximately 250 patients had follow-up visits at 24 months.

Poolability:

The Breslow-Day test was used to assess the homogeneity of NDI, neurological, FSU, and overall success results across the sites.

Covariate Adjustment:

The sponsor did not provide details on how the effects of covariates were adjusted in the primary analysis. These covariates may include, but are not limited to age, gender, weight, race, preoperative medication usage, and preoperative clinical endpoints.

PATIENT ACCOUNTING AND DEMOGRAPHICS

A total of 541 patients participated (276 investigational and 265 control patients) at 34 sites. Thirty-six (36) patients in the investigational group and 48 in the control group declined participation prior to surgery.

Patient Accountability based on Overall Success

	12 Months		24 Months	
	Invest.	Control	Invest.	Control
Enrolled	276	265	276	265
Theoretical FU	276	265	137	148
# Expected	276	263	137	148
Pts. Overall Success w/o FSU (% of Total Enrolled)	263 (95.3%)	223 (84.8%)	128 (46.4%)	122 (46.1%)
Pts. Overall Success w/ FSU (% of Total Enrolled)	205 (74.3%)	173 (65.8%)	95 (34.4%)	90 (34%)

FDA Question for Panel:

Efficacy evaluations were performed on the first 250 patients (128 Investigational, 122 Control) that had complete overall success outcome information (without FSU). This represents about 46% of the total patient enrollment in the study. In addition, even fewer patients (95 investigational and 90 control, about 34%) had complete overall success outcome information with FSU. Please comment on the appropriateness of making study conclusions using this interim analysis based on the overall success criteria with and without FSU, i.e., 46% and 34% of patients.

Reasons for Declination Prior to Surgery

	Investigational	Control
Insurance Denied	11	10
Condition Improved	5	8
Dissatisfied with Randomization	2	11
Inclusion/Exclusion Criteria Not Met	7	4
Combination*	0	1
Other**	7	12
Unknown	4	2
Total	36	48

*Combination of Condition Improved and Dissatisfaction with Randomization

** "Other" includes the following needed 2-level ACDF decided not to participate, went to another surgeon, no-show for surgery, wanted larger settlement/seeking new lawyer, posterior/lateral approach required, waiting on attorney to approve surgery, EMG and nerve conduction study indicated carpal tunnel syndrome and no radiculopathy, "got cold feet" and not cleared for surgery

Demographics:

The following tables provide summary and comparisons of demographic variables, preoperative medical condition and medication usage, preoperative evaluation of clinical endpoints, and surgery and discharge information between the investigational and control groups.

The investigational and control groups are comparable in demographic and baseline characteristics, except for alcohol use which is less than 5%.

Demographic Information

Variables	Investigational (N=276)	Control (N=265)	p-value
Age (years)	43.3 ± 7.6	43.9 ± 8.8	0.435
Height (inches)	67.4 ± 3.9	67.5 ± 4.2	0.767
Weight (lbs.)	181.7 ± 39.7	184.7 ± 41.5	0.389
Sex (% male)	46.4%	46.0%	1.000
Race			
Caucasian	260	243	0.448
Black	6	13	
Asian	1	2	
Hispanic	7	6	
Other	2	1	
Marital Status			
Single	44	32	0.240
Married	188	204	
Divorced	36	24	
Separated	5	3	
Widowed	3	2	
Education Level			
< High School	10	14	0.458
High School	73	77	
> High School	193	173	
Worker's Compensation	11.6%	13.2%	0.603
Unresolved Spinal Litigation	10.9%	12.1%	0.687
Tobacco Used	34.4%	34.7%	1.000
Alcohol Used	43.5%	53.2%	0.025
Preoperative Work Status	65.9%	62.6%	0.473

Preoperative Medical Condition and Medication Usage

Variables	Investigational (N=276)	Control (N=265)	p-value
Time to have symptoms leading to planned surgery			
< 6 weeks	21	15	0.435
6 weeks to 6 months	81	89	
> 6 months	174	161	
Number of previous neck surgeries			
0	275	263	0.745
1	1	1	
2	0	1	
Non-Narcotic medications	197 (71.9%)	187 (71.1%)	0.849
Weak Narcotic medications	130 (47.3%)	127 (48.3%)	0.863
Strong Narcotic medications	57 (20.9%)	58 (22.0%)	0.833
Muscle Relaxant medications	119 (43.4%)	114 (43.2%)	1.000

Preoperative Evaluation of Clinical Endpoints

Variables	Investigational (N=276)	Control (N=265)	p-value
NDI	55.7 ± 14.8	56.4 ± 15.9	0.632
SF-36 PCS	31.9 ± 7.0	32.0 ± 7.5	0.760
SF-36 MCS	42.4 ± 12.1	42.7 ± 12.4	0.795
Neck Pain Score	68.2 ± 22.7	69.3 ± 21.5	0.553
Arm Pain Score	59.1 ± 29.4	62.4 ± 28.5	0.191

CLINICAL STUDY RESULTS

Surgical Results and Hospitalization

The mean operative times and mean hospitalization times were statistically different for the investigational and control groups, but the difference was not clinically significant.

Surgical Results

Variables	Investigational (N=276)	Control (N=265)
Mean Operative Time (hrs)	1.6 ± 0.6	1.4 ± 0.5
Mean EBL (ml)	60.1 ± 60.3	57.5 ± 68.1
Hospitalization (days)	1.1 ± 0.6	1.0 ± 0.5
Spinal Level Treated		
C3-C4 (%)	7 (2.5)	10 (3.8)
C4-C5 (%)	14 (5.1)	15 (5.7)
C5-C6 (%)	143 (51.8)	149 (56.2)
C6-C7 (%)	112 (40.6)	91 (34.3)
External Orthosis		
Soft Collar (%)	84 (30.4)	109 (41.4)
Hard Collar (%)	1 (0.4)	40 (15.2)
None (%)	190 (68.8)	108 (40.9)
Other (%)	1 (0.4)	7 (2.7)
Operative Approach		
Extrapharyngeal Anterolateral (%)	276 (100)	265 (100)
Patient Classified as		
Inpatient (>23 hrs stay)	194	183
Outpatient (≤23 hrs stay)	82	82

Over 90% of the patients in both groups had procedures at either C5-6 or C6-7.

The distribution of treatment levels and use of orthosis postoperatively were comparable between the two treatment groups.

EFFECTIVENESS EVALUATION

The interim statistical data analysis demonstrates that the clinical results for the Prestige group were non-inferior to the control group for the clinical and radiographic endpoints. Study success was expressed as the number of individual subjects categorized as a success divided by the total number of subjects evaluated. It should be noted that outcomes for patients diagnosed with myelopathy were pooled with the outcomes for patients with radiculopathy as a diagnosis. The table below describes the success rates for individual outcome parameters and overall success. All success rates were based on the data from the 24-month follow-up evaluation and posterior probabilities of success were calculated using Bayesian statistical methods. The conclusions were based on an interim analysis which was pre-defined in the protocol.

**Posterior Probabilities of Success at 24 Months
With the First 250 Evaluable Patients**

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
NDI	80.8% (74.7%, 87.0%) (n=128)	80.8% (74.1%, 86.7%) (n=122)	0.0% (-8.9%, 8.7%)	98.5%	50.0%
Neurological	92.1% (87.6%, 96.2%) (n=128)	84.7% (78.6%, 90.5%) (n=122)	7.3% (-0.1%, 15.0%)	100%	97.1%
FSU Height	95.4% (91.5%, 98.7%) (n=128)	93.7% (89.2%, 97.8%) (n=122)	1.7% (-4.3%, 7.4%)	100%	71.7%
Overall Success without FSU	78.8% (72.1%, 85.0%) (n=128)	70.0% (62.7%, 77.4%) (n=122)	8.8% (0.9%, 18.7%)	100%	95.9%
Overall Success with FSU	80.1% (73.1%, 87.4%) (n=95)	64.0% (55.3%, 72.8%) (n=90)	16.0% (4.9%, 27.9%)	100%	99.7%

The primary endpoint for the clinical investigation was a composite variable termed “overall success.” An alternate overall success assessment was made using functional spinal unit (FSU) height maintenance or improvement along with the aforementioned criteria. Investigational treatment success was based on the 24-month overall success rate being statistically non-inferior to the control group rate.

The sponsor also performed statistical analyses on “all currently available data” as of May 9, 2006, i.e. 223 Prestige and 198 control subjects. The analysis of these patients yielded similar results to the interim analysis above (i.e., non-inferiority was demonstrated for each of the individual and composite endpoints). However, this PMA is based on the interim analysis performed on 137 investigational and 148 subjects who had reached their 24 month follow-up visit.

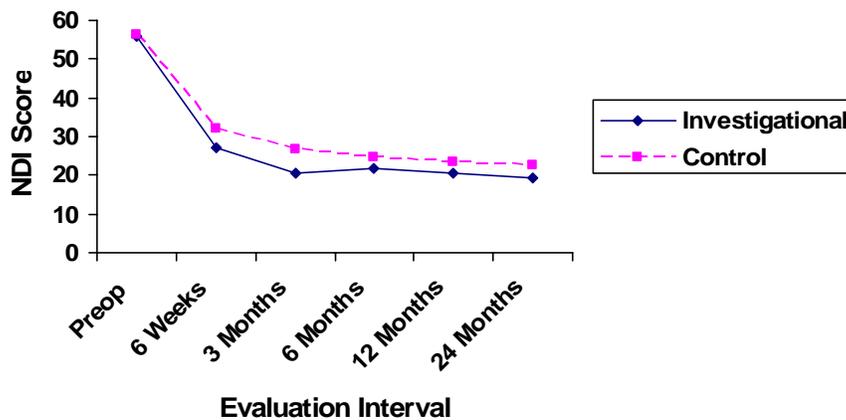
The following table summarizes the results of the Bayesian analysis of the clinical outcomes studies.

Primary Endpoints	Demonstration of Non-Inferiority
NDI	Yes
Neurological	Yes
FSU Height	Yes
Overall Success (without FSU)	Yes
Overall Success (with FSU)	Yes
Secondary Endpoints	
Radiographic Success	Yes
SF-36 PCS	Yes
SF-36 MCS	Yes
Neck Pain	Yes
Arm Pain	Yes
Doctor's Perception	Yes
Patient Satisfaction	Yes
Patient Perceived Global Effect	Yes
Gait Assessment	Yes
Foraminal Compression Test	Yes

PRIMARY OUTCOME VARIABLES

NDI

The NDI success criterion is a function of the preoperative NDI score. A 15-point or greater NDI score improvement following surgery was required to be deemed a successful outcome. At all postoperative time periods for both treatment groups, the mean overall NDI scores improved when compared to the preoperative scores, and these improvements were highly statistically significant ($p < 0.001$). The mean improvements in NDI scores for the investigational group at 12 and 24 months postoperative were 34.8 and 35.2, respectively; these values are greater than the mean improvement scores of 32.8 and 33.6 for the control group and greater than the 15 point difference which is considered clinically significant.



NDI score improvement following surgery exceeded 30 points at 12 and 24 months for both groups, showing non-inferiority.

Neck Disability Index	Investigational (n=128)	Control (n=121)
Success	106 (82.8%)	99 (81.8%)
Failure	22 (17.2%)	22 (18.2%)

Bayesian statistical analyses showed that the posterior probability of non-inferiority of investigational group to the control is 98.5%, thus demonstrating statistical non-inferiority.

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
NDI	80.8% (74.7%, 87.0%)	80.8% (74.1%, 86.7%)	0.0% (-8.9%, 8.7%)	98.5%	50.0%

The numbers of this two tables are different (82.8% vs. 80.8%) because the first one is simple a mean proportion (106/128) and the second one is a Bayesian posterior mean obtained by combining 128 complete patients and all other patients with only 12 month evaluation. The differences in the tables that follow are due to the same reason.

Neurological

Neurological success for each of the three indicators was based on maintenance or improvement of condition postoperatively as compared to the preoperative status for each element. Overall neurological status success was based on demonstrating maintenance or improvement in all three indicators.

	Investigational (n=128)	Control (n=121)
Motor		
Success	126 (98.4%)	117 (96.7%)
Failure	2 (1.6%)	4 (3.3%)
Sensory		
Success	123 (96.1%)	111 (91.7%)
Failure	5 (3.9%)	10 (8.3%)
Reflexes		
Success	127 (99.2%)	115 (95.0%)
Failure	1 (0.8%)	6 (5.0%)
Overall		
Success	120 (93.8%)	105 (86.8%)
Failure	8 (6.3%)	16 (13.2%)

The neurological success rates at 24 months postoperative were 93.8% and 86.8% for the investigational and control groups, respectively.

Bayesian statistical analyses showed that the posterior probability of non-inferiority of investigational group to the control is 100%, thus demonstrating statistical non-inferiority.

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
Neurological	92.1% (87.6%, 96.2%)	84.7% (78.6%, 90.5%)	7.3% (-0.1%, 15.0%)	100%	97.1%

FSU

Despite using FSU height as an indicator of disc space height maintenance, according to the sponsor, measurements were encumbered by the inability to visualize the area of interest or a poor-quality film. This was especially true for patients having procedures at C6-C7, where the

shoulders can obscure the area of interest. In addition, both 6-week and 24-month measurements were needed to obtain FSU success; if either was missing, the FSU success result will be missing. This parameter was part of the original and approved investigational study design. For the interim analysis cohort, success/failure determinations could be made for approximately 70% of the patients who had 24 month data, i.e. 95 investigational and 90 controls, and approximately two-thirds of the missing FSU results were in patients having C6-C7 procedures. The missing FSU success results were spread fairly evenly between the investigational and control groups.

For the available data, the rates of FSU height maintenance or improvement at 3, 6, 12 and 24 months following surgery exceeded 95%, for both treatment groups at the four postoperative time periods. Bayesian analyses comparing the overall investigational FSU height success rate to that for the control group demonstrated a posterior probability of non-inferiority value of essentially 100%, thereby demonstrating statistical non-inferiority.

Functional Spine Unit	Investigational (n=94)	Control (n=88)
Success	91 (96.8)	84 (95.5)
Failure	3 (3.2)	4 (4.5)

Bayesian statistical analyses showed that the posterior probability of non-inferiority of investigational group to the control is 100%, thus demonstrating statistical non-inferiority.

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
FSU Height	95.4% (91.5%, 98.7%)	93.7% (89.2%, 97.8%)	1.7% (-4.3%, 7.4%)	100%	71.7%

Overall Success

24 Month Overall Success Rate

	Investigational (n=128)	Control (n=121)
NDI		
Success	106 (82.8)	99 (81.8)
Failure	22 (17.2)	22 (18.2)
Neurological		
Success	120 (93.8)	105 (86.8)
Failure	8 (6.3)	16 (13.2)
FSU		
Success	91 (96.8)	84 (95.5)
Failure	3 (3.2)	4 (4.5)
Second Surgery Failure	3	6
SAE	4	5
Overall Success w/out FSU		
Success	103 (80.5)	87 (71.3)
Failure	25 (19.5)	35 (28.7)
Overall Success w/ FSU		
Success	77 (81.1)	58 (64.4)
Failure	18 (18.9)	32 (35.6)

At 24 months following surgery, the overall success rate for the investigational group was 80.5%, as compared to a 71.3% rate for the control group. Bayesian statistical analyses yielded a posterior probability of non-inferiority at 24 months of essentially 100%. The posterior probability of superiority was found to be 95.9%.

Overall success rates were also calculated using the same criteria mentioned above with the addition of FSU (disc height) success. FSU success status could be determined only in approximately 75% of the patients in the interim analysis cohort who had NDI or neurological status data (two of the overall success components). The missing FSU success results were spread fairly evenly between the two treatment groups.

In this study, for the available data, FSU success rates exceeded 95%. A few failures in both groups are possibly due to measurement variation in combination with the success criteria used. According to the sponsor, “supporting this argument is the fact that the investigational group had slightly higher success rates at both 12- and 24-month evaluations than the control group, in which the treatment is an instrumented fusion and any subsidence (FSU failure) is unusual. Thus, essentially every patient in both groups has a successful FSU status when data were available.”

At 24 months following surgery, the overall success rate, including FSU for the investigational group was nearly 17 percentage points higher than the control group (81.1% vs. 64.4%). Despite the smaller sample size due to missing FSU success values, Bayesian statistical analyses of these modified overall success rates yielded a posterior probability of non-inferiority at 24 months of essentially 100%. The posterior probability of superiority was found to be 99.7%.

Secondary Endpoints:

Neck Pain

Success was determined by comparing the postoperative composite neck pain score to the preoperative score on a patient basis. Success was based on the patient having no worsening in neck pain score following surgery. At 12 and 24 months postoperative, the investigational group had neck pain success rates of 94.7% and 93.8%, respectively. The control group rates were 95.5% and 99.2%, respectively.

Neck Pain	Investigational (n=128)	Control (n=121)
Success	120 (93.8)	120 (99.2)
Failure	8 (6.2)	1 (0.8)

The Bayesian statistical analyses showed that the posterior probability of non-inferiority of the investigational device to the control at 24 months is 99.2%, i.e., statistically non-inferior.

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
Neck Pain	93.6% (89.9%, 97.2%)	97.7% (95.1%, 99.8%)	-4.0% (-8.5%, 0.5%)	99.2%	3.7%

Arm Pain

Success was based on the patient having no worsening in arm pain score following surgery. At 12 months postoperative, the arm pain success rate for the investigational device group was 89.4%, as compared to a 92.3% rate for the control group. At 24 months, the rates were 90.6% and 94.2%, respectively.

Arm Pain	Investigational (n=128)	Control (n=121)
Success	116 (90.6)	114 (94.2)
Failure	12 (9.4)	7 (5.8)

Bayesian statistical analyses showed that the posterior probability of non-inferiority of the investigational group to the control is 98.1%, thus demonstrating non-inferiority.

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
Arm Pain	90.0% (85.2%, 94.6%)	93.0% (88.8%,96.9%)	-3.0% (-9.7%, 3.2%)	98.1%	17.3%

Patients whose Neck or Arm Pain Increased from Baseline

	Investigational (n=225)	Control (n=197)
Neck Pain	10	3
Arm Pain	17	11

Ten patients in the investigational group and three in the control group had neck pain scores at last follow-up that were greater than their scores at baseline. Similarly, 17 patients in the investigational group and 11 in the control group had arm pain scores at last follow-up that were greater than their scores at baseline. Although these events are small in number, FDA is unsure of the significance of these events, particularly since they occur in greater frequency in the investigational group.

SF-36

For the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) higher scores represent higher levels of health.

Variable	24 month outcome	Investigational (N=128)	Control (N=122)
PCS	Success	109 (85.8)	102 (85.7)
	Failure	18 (14.2)	17 (14.3)
MCS	Success	84 (66.1)	88 (73.9)
	Failure	43 (33.9)	31 (26.1)
Physical Function	Success	109 (85.2)	106 (88.3)
	Failure	19 (14.8)	14 (11.7)
Role-Physical	Success	117 (92.1)	116 (97.5)
	Failure	10 (7.9)	3 (2.5)
Pain Index	Success	118 (92.2)	115 (95.8)
	Failure	10 (7.8)	5 (4.2)
General Health Perception	Success	74 (57.8)	67 (55.8)
	Failure	54 (42.2)	53 (44.2)
Social Function	Success	108 (84.4)	103 (85.8)
	Failure	20 (15.6)	17 (14.2)
Mental Health	Success	89 (69.5)	94 (78.3)
	Failure	39 (30.5)	26 (21.7)
Role-Emotional	Success	112 (87.5)	111 (92.5)
	Failure	16 (12.5)	9 (7.5)
Vitality	Success	108 (84.4)	97 (80.8)
	Failure	20 (15.6)	23 (19.2)

Bayesian statistical analyses showed that the posterior probability of non-inferiority of the investigational group to the control group is 97.9% for the PCS component and 87.5% for the

MCS component. Thus, non-inferiority is demonstrated for the PCS component, but not the MCS component.

Primary Outcomes	Investigational Pt	Control Pc	Difference Pt - Pc	Non-Inferiority	Superiority
PCS	82.8% (76.5%, 88.9%)	83.8% (77.2%, 89.9%)	0.9% (-7.8%, 9.9%)	97.9%	41.7%
MCS	66.9% (59.6%, 73.9%)	70.8% (63.7%, 78.4%)	3.9% (-6.6%, 13.9%)	87.5%	23%

All mean PCS and MCS postoperative scores were higher than preoperative scores for both treatment groups. The mean improvement in PCS scores from preoperative to 12 and 24 months following surgery for the investigational group (12.8 and 12.9) compared very favorably to those values for the control group (11.2 and 11.4, respectively). The mean improvements in MCS scores from preoperative to 12 and 24 months postoperative for the investigational patients (7.7 and 7.1) were also comparable to those values for the control group (6.1 and 8.5).

Success was defined as the proportion of patients who demonstrated maintenance or improvement in SF-36 results postoperatively as compared to the preoperative condition. For the PCS results at 12 and 24 months, the success rates for the two treatment groups exceeded 85% and were very similar. For the MCS, the investigational group rate at 12 months exceeded that of the control group by nearly eight percentage points. However, at 24 months, the success rates reversed with the control group rate being almost eight percentage points higher.

Bayesian analyses were performed comparing both the 24-month PCS and MCS results of the investigational group to the control group. For the PCS results, the posterior probability of non-inferiority was found to be 97.9%. The posterior probability of non-inferiority comparing the MCS success results of the investigational group to the control group was 87.5%. Therefore, statistical non-inferiority was demonstrated for the PCS comparison, but not for MCS. However, the mean MCS improvement scores (preoperative vs. 24 months) were 7.1 and 8.5 points for the investigational and control groups, respectively. This small difference of 1.4 points was not statistically different ($p=0.480$, t-test).

Global Perceived Effect

At each postoperative time period, patients were asked to evaluate their overall impression of their study treatment effectiveness as a function of pain. The seven possible answers ranged from “completely recovered” to “vastly worsened”. At 12 and 24 months following surgery, 81.0% and 85.1%, respectively, of the investigational patients indicated that they had either “completely recovered” or were “much improved”. These rates were higher than the 74.9% and 81.0% rates, respectively, for the control group.

Doctor’s Perception of Results

At each postoperative visit, the doctors were asked to provide their perceptions of the patients’ conditions. The responses could be “excellent”, “good”, “fair”, or “poor”. At 12 months following surgery, 90.9% of the doctors responded that investigational patients were in “excellent” or “good” condition. This rate is higher than the 87.5% value for the control group. At 24 months postoperative, 94.5% of the investigational device and 91.7% of the control responses were either “excellent” or “good”. These findings show that a substantial majority of patients in both treatment groups were progressing well clinically in the overall opinions of the doctors.

Treated Level Measurements

Investigational Group

Angular motion was measured at each study period by comparing lateral flexion and extension radiographs. The pre operative level of motion was maintained following the implantation of the investigational device.

	Pre Operative	12 months	24 months
Angular motion (mean)	7.55°	7.59°	7.87°
Translational motion (mean)	0.26 mm	0.33 mm	0.28 mm

Translational motion was also measured throughout the course of the study by comparing lateral flexion and extension radiographs. Again, the postoperative values approximated the preoperative determinations.

Radiographic Success

Investigational Patients	Angular Motion >4° to ≤20°	No Bridging Bone	Overall Radiographic Success
24 Months Success (%)	85 (73.3)	122 (99.2)	85 (72.6)
Failure (%)	31 (26.7)	1 (0.8)	32 (27.4)

At 12 and 24 months following surgery, the radiographic success rates were 69.9% and 72.6%, respectively. The primary contributor to these success rates was the angular motion component since bridging bone was not observed in many patients – only one at 24 months. The angular motion component yielded success rates between 70 and 77% at the postoperative time periods. The mean values were consistently in a range from 7 to 8 degrees.

Lateral bending was evaluated by comparing the angular movements from left and right neck bending films. Throughout the postoperative course, the mean results were very consistent in a range of 6.36° to 6.80°.

Control Group

Radiographic success for control patients at 12 and 24 months following surgery, were 98.7% and 98.8%, respectively.

FDA Question for Panel:

The package insert currently has no claims of the device maintaining range of motion. The sponsor does state in the PMA, however, that in the treatment group, the mean preoperative angular motion (flexion/extension) at the target segment was 7.55° while the mean angular motion values at 12 and 24 months were 7.59° and 7.87°, respectively. Is this data adequate to demonstrate a possible claim that the device maintains motion? Please discuss how the labeling might be adjusted to reflect such information.

Adjacent Level Measurements
Angular Motion

	Pre-op		12 months		24 months	
	Invest.	Control	Invest.	Control	Invest.	Control
Level above treated segment (Mean)	11.17°	10.77°	11.94°	12.07°	12.05°	11.63°
Level below treated segment (Mean)	8.32°	7.77°	8.33°	9.53°	9.47°	9.07°

Motion at the level above the treated level tended to be higher than the level below the treated level. For investigational patients, both the mean values for the level above and below were higher than the angular motion value for the treated segment at 24 months (12.05° vs. 9.47° vs. 7.87°, respectively).

Translational Motion

The mean values for both treatment groups were very similar and remained fairly constant over time. For the level above the treated segment, the mean values from preoperative through 24 months following surgery were in a range of 1.31 mm to 1.27 mm for the investigational and control groups.

The translational motion values for the level below the treated segment were consistently lower than the level above the treated level by approximately 0.50 mm. For the level below the treated segment, the mean values from preoperative through 24 months following surgery were in a range of 0.92mm to 0.76 mm for the investigational and control groups.

For investigational patients at 24 months postoperative, both the level above and below had higher translational motion values than the 0.28 mm treated segment value.

Thus it appears that the levels adjacent to the treated cervical segment were stable with regard to translation movement over the postoperative course, and the motion levels were similar to those before treatment.

Gait Assessment

Patients with a normal gait without nerve root or spinal cord symptoms were classified as “normal.” 79.7% of the investigational and 76.9% of the control patients had normal gait preoperatively. Postoperatively, 99.2% of the investigational patients and 98.3% of the control patients had “normal” values at 24 months following surgery.

Foraminal Compression Test

Preoperatively, 47.6% of the investigational patients and 46% of the controls had “negative” responses. At 24 months, the rates of “negative” outcomes for both treatment groups were in excess of 95% for both groups.

Work Status

Preoperatively, approximately 66% of the investigational patients were working, as compared to a 63% rate for control patients. At 24 months following surgery, the percent of working patients in the investigational group was 78.1% and the percent of the control group was 71.9%.

A better way to examine work status is to analyze the number of days from surgery to work return using Kaplan-Meier life table methods. The median return to work value for investigational

patients was 45 days, as compared to 61 days for control patients. This 16-day difference approached statistical significance (Log-Rank Test $p=0.094$, Wilcoxon Test $p=0.022$).

Patient Satisfaction

The patients responded to the following statements:

1. I am satisfied with the results of my surgery.
2. I was helped as much as I thought I would be with my surgery.
3. All things considered I would have the surgery again for the same condition.

At 24 months postoperative for the first question, 89.0% of the investigational patients and 90.1% of the control patients responded either “definitely true” or “mostly true”. For the second question, 85.0% of the investigational patients and 85.1% of the control patients thought that they were helped as much as expected from their surgeries. Finally, 87.4% of the investigational patients said that they would have the surgery again, as opposed to an 84.3% rate for the control group.

ADDITIONAL DATA PRESENTATIONS

Examination of Effectiveness Variables by Investigator and Justification for Pooling Data Across Investigational Sites

The Breslow-Day test was used to assess the homogeneity of NDI ($p=0.119$), Neurological ($p=0.667$), FSU ($p=0.240$), and overall success ($p=0.165$) results across the sites. There were no statistically significant differences noted in any of the comparisons. However, the Breslow-Day test uses the normal assumption. Even after the sponsor has combined sites with fewer than 10 enrolled patients into one site, there are still 12 sites with less than 10 evaluable patients. Hence this test may lack power.

“Per Protocol” Results

The “per protocol” dataset was a subset of patients who were included in the primary analysis dataset. Patients who were excluded from the “per protocol” analysis had major protocol deviations, i.e., did not meet the inclusion/exclusion criteria or received the wrong study treatment, or other major protocol deviations that could potentially affect clinical outcomes.

The following table summarizes the results at 24 months following surgery.

“Per Protocol” Success Rates			
	Investigational	Control	Post. Prob. of Non-inferiority
NDI	82.5% (104/126)	83.0% (93/112)	97.1%
Neurological	93.7% (118/126)	86.6% (97/112)	100.0%
FSU Height	97.8% (90/92)	95.1% (78/82)	100.0%
Overall Success (without FSU)	80.2% (101/126)	72.6% (82/113)	100.0%
Overall Success (with FSU)	81.7% (76/93)	65.5% (55/84)	100.0%

Like the previous analyses, every statistical comparison for the “per protocol” dataset yielded a posterior probability of non-inferiority of at least 95%.

In addition, the “per protocol” dataset was further refined by excluding any “out of window” visits and similar analyses were performed on it. The statistical analyses showed that the results and conclusions were very similar to those obtained with “out of window” data included, and investigational group overall success outcomes were still non-inferior to those of the control group.

“Missing Equals Failure” Results

The “missing-equals-failure” data presentations for various study periods are included in Attachment Q. For this presentation, secondary surgery failures, deaths, patients lost-to-follow-up, and missing observations due to other causes resulted in missing observations for the outcome variables and, therefore, were included in the denominators of the calculated rates, i.e., considered as “failures.” By treating these patients as treatment failures, the clinical outcome rates in the “missing-equals-failure” analyses were lower than those observed in the clinical data.

The 24-month overall success rate (without FSU) for the investigational group was higher than that of the control group (75.2% vs. 58.8%).

The same is also true for the investigational group with FSU in the definition (56.2% vs. 39.2%).

Missing Equals Failure Success Rates			
Variable	24 month outcome	Investigational (N=137)	Control (N=148)
NDI	Success	106 (77.4)	99 (66.9)
	Failure	31 (22.6)	49 (33.1)
Neurological	Success	120 (87.6)	105 (70.9)
	Failure	17 (12.4)	43 (29.1)
FSU	Success	91 (66.4)	84 (56.8)
	Failure	46 (33.6)	64 (43.2)
Overall Success without FSU	Success	103 (75.2)	87 (58.8)
	Failure	34 (24.8)	61 (41.2)
Overall Success with FSU	Success	77 (56.2)	58 (39.2)
	Failure	60 (43.8)	90 (60.8)

Sensitivity Analysis for Assessing Missing Values

It should be noted that there was a disparity in follow-up rates at 24 months between the investigational and control group. In the interim analysis cohort, nine (6.6%) of 137 investigational patients did not have overall success outcomes, as compared to 26 (17.6%) of 148 control patients. To assess the impact of lost-to-follow-up on study conclusions, a sensitivity analysis was performed of overall success at 24 months by various imputations for the missing outcomes. The analyses were focused on the 24-month data and used simple frequentist calculations.

The results show that even in the worst case scenario (where all missing investigational outcomes are assumed to be failures and all missing control outcomes are assumed to be successes), which is grossly biased against the investigational group, non-inferiority of the investigational treatment to the control can still be claimed (p=0.0411). When 50% of missing investigational outcomes and 60% of the missing control outcomes are assumed to be successes (which favors the control group and could perhaps be closer to the actual situation), the superiority of the investigational treatment to the control can still be claimed (p=0.0363).

Correlations between 12-Month and 24-Month Results

Analyses were performed to examine the relationships between certain key endpoints at 12 and 24 months postoperative. The results for the primary and “per protocol” dataset are presented in the table below.

Percent Agreement Between 12- and 24-Month Data				
	Primary Dataset		“Per Protocol” Dataset	
	Invest. (n=128)	Control (n=121)	Invest. (n=126)	Control (n=112)
NDI	88.1%	87.0%	88.7%	86.9%
Neurological	92.9%	88.8%	92.7%	88.0%
FSU Height	97.8%	97.5%	98.9%	98.7%
Overall Success (without FSU)	84.9%	83.6%	85.5%	83.3%
Overall Success (with FSU)	84.4%	80.7%	86.4%	80.8%

There is good agreement between the 12- and 24-month outcomes. This means that there is a high likelihood of a patient in either treatment group having the same outcome at the two latter study periods. This is especially important for Bayesian analyses since it strengthens the inferences that can be made.

Correlation between Pain and Disability Outcomes and Angular Motion Measurements

The relationships between NDI, neck pain, and arm pain results and angular motion values were examined in investigational device patients to determine if there was any correlation between the degree of segmental motion and pain. At 6 weeks following surgery, there was no significant correlation between these measurements. However, at later postoperative intervals, statistically significant correlations were noted. At 12 and 24 months following surgery, the three indicators of pain, i.e., NDI, neck pain, and arm pain, were found to be negatively correlated to angular motion, and in all comparisons, the correlations were statistically significant, although the magnitude of the correlations was very moderate.

Financial Disclosure Information and Analyses

20 of 72 (28%) surgeons who performed surgeries met the criteria for having a financial interest at some point during the course of the clinical study. These surgeons contributed 187 patients to both treatment groups. According to the sponsor’s calculations, at 12 and 24 months postoperative, there were no statistically significant differences in any of the outcome comparisons between the patients of surgeons with a financial interest versus those without.

FDA Question for Panel:

Please discuss whether the clinical data in the PMA provide reasonable assurance that the proposed device is effective. If not, what additional data or analyses are needed?

SAFETY EVALUATION

The safety of the investigational device was evaluated based on the nature and frequency of adverse events, as compared to those occurring in the control group. The adverse effects, as shown in the table below, were reported from the 276 PRESTIGE® device patients and 265 control patients enrolled in the multi-center clinical study. A total of 226 (81.9%) investigational patients had at least one adverse event, as compared to 212 (80.0%) for the control group. The number of patients having serious adverse events, i.e. those with a World Health Organization (WHO) grade of 3 (severe) or 4 (life threatening), in the investigational group was 77 (27.9%), as compared to 79 (29.8%) in the control group. The number of patients having adverse events that are classified as implant-associated in the investigational group was 9 (3.3%), as compared to 26 (9.8%) in the control group.

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.

Adverse Event	Investigational (N patients, % 276, N events)	Control (N patients, % 265, N events)
Any Adverse Event	226 (81.9)	212 (80.0)
Anatomical/Technical Difficulty	1 (0.4) 1	0 (0.0) 0
Cancer	5 (1.8) 5	2 (0.8) 2
Cardiovascular	14 (5.1) 15	8 (3.0) 9
Carpal Tunnel Syndrome	12 (4.3) 14	7 (2.6) 7
Death	0 (0.0) 0	3 (1.1) 3
Dysphagia/Dysphonia	23 (8.3) 23	22 (8.3) 22
Gastrointestinal	25 (9.1) 26	24 (9.1) 26
Implant Displacement/ Loosening	2 (0.7) 2	4 (1.5) 4
Infection	27 (9.8) 30	20 (7.5) 22
Neck and/or Arm Pain	138 (50.0) 190	127 (47.9) 173
Neurological	66 (23.9) 78	55 (20.8) 65
Non-Union	0 (0.0) 0	6 (2.3) 6
Other	70 (25.4) 109	66 (24.9) 82
Other Pain	69 (25.0) 88	56 (21.1) 68
Pending Non-Union	0 (0.0) 0	16 (6.0) 16
Respiratory	8 (2.9) 8	8 (3.0) 9
Spinal Event	17 (6.2) 18	30 (11.3) 32
Subsidence	1 (0.4) 1	0 (0.0) 0
Trauma	59 (21.4) 69	40 (15.1) 47
Urogenital	15 (5.4) 16	5 (1.9) 6
Vascular Intra-Op	5 (1.8) 5	2 (0.8) 2

The rate of investigational patients having at least one AE was very similar to the control group. This was also true for serious adverse events (SAE). Investigational patients had a lower number of adverse events that were classified as implant- or implant/surgical procedure-associated. The radiographic reviewers did not note any implant migration or fractured/broken implants in the investigational group, while there were instances in the control group. The investigational group had statistically lower rates of second procedures related to revisions and supplemental fixations. The rate of removals was also lower but not statistically significant. The investigational group neurological success rate was statistically higher than the control group.

The reported rates of several adverse events were greater than 10% in both the investigational and control groups. These events included neck and/or arm pain, neurological, other, other pain, and trauma. Spinal events, at any level, occurred in 11.3% of the control patients and 6.2% of the investigational patients.

The majority of the adverse events occurred peri-operatively. However, in the investigational group there were 2 reports of cancer at the 12 month follow-up and 3 reports at the 24 month follow-up. Regarding implant displacement/loosening, there was 1 report at the 3 month visit and 1 report at the 24 month visit. The one case of subsidence was reported at 6 months.

Adverse Events Related to Device (Definite, Probable, Possible)

Adverse Event	Investigational (N pts, % 276)	Control(N pts, % 265)
Pt with any AE	9 (3.3)	26 (9.8)
Anatomical/Technical Difficulty	1 (0.4)	0 (0.0)
Implant Displacement/Loosening	2 (0.7)	3 (1.1)
Infection	0 (0)	1 (0.4)
Neck and/or Arm Pain	1 (0.4)	2 (0.8)
Neurological	4 (1.4)	1 (0.4)
Non-Union	0 (0.0)	6 (2.3)
Pending Non-Union	0 (0.0)	16 (6.0)
Subsidence	1 (0.4)	0 (0.0)

The number of adverse events that were considered to be implant- or implant/surgical procedure-associated, including implant displacement/loosening and neck and/or arm pain, were greater in the control group compared to the investigational group. However, the rates of all these events were low in both groups. Six serious (WHO Grade 3 or 4), implant- or implant/surgical procedure-associated adverse events were reported; all of these occurred in control patients. No deaths were reported among investigational patients. Three control group deaths were reported, all of which were due to myocardial infarction or cardiac arrest.

Secondary Surgical Interventions

Some of the reported adverse events required surgical interventions subsequent to the initial surgery. The number of subjects requiring a second surgical intervention classified as a revision, removal, re-operation, or supplemental fixation was 3.3% (9/276) in the investigational group and 9.1% (24/265) in the control group. The investigational group had a statistically lower rate of revisions and supplemental fixations than the control group. Investigational patients also experienced a lower rate of implant removals, but it was not statistically different. These findings resulted in a lower second surgery failure rate for investigational patients.

Secondary Surgical Procedures

	# Pts \leq 24 Months	
	Invest. (N=276)	Control (N=265)
Revisions	0	5 (1.9)
Removals*	5 (1.8)	9 (3.5)
Re-operations	4 (1.4)	2 (0.8)
Supplemental Fixations	0	8 (3.0)
Other	58 (21.0)	44 (16.6)

*This includes both elective and nonelective removals in the control group.

This table summarizes the revision, re-operations and removals for the investigational patients.

Pt. Invest.	Time	Intervention	Surgery
210	12 mo.	Re-operation	C5-6 posterior cervical laminectomy
904	12 mo.	Removal	Explant ACDF (Stryker Plate) and removal osteophytes
904	12 mo.	Re-operation	C5-6 posterior foraminotomy/discectomy and nerve root exploration
1002	12 mo.	Removal	Explant followed by ACF C6-7
1006	6 mo.	Removal	Explant followed by ACDF C6-7
1916	12 mo.	Removal	Explant ACDF C5-7
2802	24 mo.	Removal	Removal
2855	12 mo.	Re-operation	Cervical Foraminotomy C5-6
2902	6 wks.	Re-operation	Decompressive Foraminotomy C4-5

This table summarizes the revisions, re-operations and removals for the control patients.

Pt. Control	Time	Intervention	Surgery
215	12 mo.	Removal/Elective	ACDF C5-6 (Fusion at C6-7 solid; Atlantis plate removed)
401	24 mo.	Supplemental Fixation	Posterior cervical fusion C5-6 for nonunion
608	12 mo.	Re-operation	Cervical laminotomy C4, 5, 6, 7 and T1
632	6 mo.	Removal	Removal Atlantis plate and allograft; revision discectomy C5-6, bilateral foraminotomies and anterior arthrodesis w iliac crest and stem bone
1112	12 mo.	Revision	Removal followed by ACDF C5-6-7
1609	Postop	Revision	Left foraminotomy C5-6; Atlantis plate removed, disc space cleared, nerve root decompressed; Atlantis plate replaced
1713	3 mo.	Revision	Discectomy and fusion C5-7; original Atlantis plate removed and replaced with 2-level Atlantis plate
1720	12 mo.	Supplemental fixation	Posterior cervical fusion C5-6 with vertex lateral mass screws and rods
2507	12 mo.	Removal	Revision arthrodesis C5-7 with iliac crest bone and Atlantis vision plate
2611	6 wks	Reoperation	Posterior C5 foraminotomy
2705	12 mo.	Removal	Removal Atlantis plate C5-6 with partial corpectomy at C6, microdissection, cervical fusion C5-6 with cornerstone graft, crushed cancellous autograft, small kit BMP and Orion plate
2716	12 mo.	Removal	Atlantis plate and allograft removed; replace with cornerstone, BMP and Orion plate
2719	6 mo.	Removal	Partial corpectomy C5, decompression and microdissection, ACDF C5-6 with cornerstone prosthetic graft/BMP/allograft and Orion plate
2805	12 mo.	Removal	Removal Atlantis plate and allograft; replace with iliac crest bone and another Atlantis plate
3117	12 mo.	Removal/elective	Cervical discectomy and fusion C6-7. Atlantis plate electively removed C5-6
3310	3 mo.	Revision	Removal Atlantis plate and placement C5-6-7 segmental instrumentation (2 level Atlantis plate). Original allograft C6-7 left intact.

There were no reported revision procedures in the investigational group. There were five revisions (1.9%) in the control group. For control patients, four of the revision procedures

involved an adjacent level fusion. The remaining revision procedure occurred shortly after the original procedure to remove residual disc material.

Likewise, no patients in the investigational device group had a supplemental fixation procedure. Eight control patients had nine procedures. Seven of these procedures were due to suspected non-unions arising from the original procedures. The other two procedures were performed in response to pain and neurological symptoms. Six of the nine reported supplemental fixations were attributed to the use of bone growth stimulators.

Implant removals occurred in both treatment groups. The removal rate was 1.8% in the investigational group and 3.4% in the control group. The removals in the investigational group were primarily due to the treatment of symptoms such as pain and neurological complaints. Fusion procedures followed these removals. Seven of the nine control implant removals were non-elective, while two were elective removals. The non-elective removals were often associated with additional fusion procedures in which different implants were used. One non-elective removal in the control group occurred shortly after the original procedure in the treatment of an esophageal abscess.

Five (5) investigational patients had implant removal procedures. Histological and metallurgical analyses for three of these cases have been performed and are discussed in the preclinical section. The histological analyses found tissue responses consistent with those typically seen in proximity to metal-on-metal arthroplasty devices. In the metallurgical analyses, most of the implant surfaces showed only superficial wear patterns, and there was no evidence of fracture or damage that would suggest a manufacturing or processing defect.

Cumulatively, the investigational group had five second surgery “failures”, as compared to 12 for the control group. One of the “failures” in the investigational group occurred after 24 months postoperative. At or before 12 months, there were two second surgery failures in the investigational group and nine in the control group. Cumulatively, the investigational group had five second surgery “failures”, as compared to 12 for the control group.

Neurological Adverse Events

A total of 78 neurological events occurred in 66 patients in the investigational group (23.9%). A total of 65 neurological events occurred in 55 patients in the control group (20.8%).

Severity Rating of Adverse Events

	Investigational	Control
Total Events/Patients	78/66	65/55
Grade 1	20	20
Grade 2	50	39
Grade 3	8	6
Grade 4	0	0

The following table summarizes the neurological events that occurred during the study. Following the table is a detailed description of each of the neurological adverse events.

Event Type	Investigational	Control
Numbness	22	18
Paresthesia	5	5
Tingling	4	3
Numbness/Tingling	7	7
Numbness/Pain	5	5
Radiculopathy	11	0
Weakness	4	5
Paresthesia/Pain	2	0
Numbness/Weakness	2	1
Numbness/Tingling/Pain	0	4
Other*	21	15

*Events that occurred in only one patient.

Investigational

Out of the 78 events, the most commonly reported event among investigational patients was numbness (22 events). Of these 22 events, 17 involved the upper extremities (arms, hands and fingers).

The next most frequently reported neurological events in investigational patients involved paresthesia, tingling, numbness and tingling, numbness and pain, and radiculopathy. There were five events of paresthesia affecting the arm, feet, and hands; two events affecting the hand; and one event of nocturnal paresthesia in both hands. There were four events of tingling affecting the hand, the elbow to the left hand, bilateral arms, and the hands and fingers. There were seven events of numbness and tingling affecting the bilateral upper extremities, left hand, and the right middle finger, and two events affecting the left arm, the hands, and the wrist. There were five events of numbness and pain affecting the hand and elbow, fingers, neck, shoulder, and back, and two events affecting the arm. There were six events of radiculopathy in the C8 area, two events affecting the arm, two events affecting the left side, and one event involving neck spasms.

There were four events of weakness and two events of paresthesia and pain. The four events of weakness affected the elbow, shoulder, or arm (two events). The two instances of paresthesia and pain involved arm/cervical radiculopathy and the left arm or neck. There were two events of numbness and weakness in the leg and foot or arms and hands.

There were 21 events that only occurred once in the investigational group. These events included: C5 radiculopathy associated with shoulder pain; paresthesias/hypothesias; numbness and paresthesias; sciatica; Bell's Palsy; axillary pain/bilateral; bodily shocking sensations; burning and tingling in the hand; radiating pain in the neck; dizziness and numbness; decreased sensation in the little finger with neck, right shoulder, and right arm pain; numbness in the hands and arms with neck spasms; restless legs; pain with "needles" in the left arm and right foot; "pins and needles" in the thumbs and forearms; left trapezius shooting pain and weakness; neck swelling and stiffness with left hand numbness; numbness and weakness in the foot with low back pain; numbness, tingling, and pain the arm; pain radiating in the right arm associated with numbness; and back and leg pain thought to be associated with a previous L5-S1 disc bulge.

Control

Out of the 65 neurological events that occurred in control patients, the most frequently reported event was numbness. There were 15 events that involved the upper extremities (arms, elbow, hands and fingers). In addition, there were two occurrences of chin numbness. There was one instance involving numbness of the bilateral lower extremities, and one report of general numbness involving the right side at C6-C7.

The next most frequently reported neurological event involved numbness accompanied by tingling, pain, tightness, and/or burning. There were four events involving numbness associated with tingling and pain. These events involved the right arm, shoulder and neck, finger and hand with shoulder pain, and thumb with neck pain. There were seven events that involved only numbness and tingling. These events included one instance involving the bilateral upper and lower extremities; two instances involving the bilateral hands; one instance involving the arms, hands, and feet; one instance involving the right arm and fingers; one instance involving the low back; and one instance involving the neck.

There were three events involving numbness associated with radiating pain. These events included one occurrence each in the groin; the arms and legs; and in the back, hip, thigh, and foot. There were two events that involved numbness and pain. These events involved numbness in the thigh with general spine pain, and numbness with pain in the shoulder and finger.

The next most commonly reported events involved paresthesia, weakness, and tingling. There were five events involving paresthesia. These were located in the thigh, calf, and foot, the left leg and foot, the right hand, the arm and hand, and the fingers of the left hand. There were five events that involved weakness of the right tricep, left arm, right arm, left deltoid, and bicep. There were three events involving tingling. These events included tingling in the right leg, the right hand, and the left hand.

Finally, there were two events that involved sciatica, of which one included numbness and weakness in the left leg. There were 15 events that only occurred once. These events included: hypersensitivity, myelopathy, hyperpathia, decreased pin prick, dysesthesia of the third and fourth digits, ulnar neuropathy, muscle hyperexcitability and twitching, hemisensory loss of the left side along with walking difficulty and visual obscuration, seizure, transverse myelitis, involuntary movements of the thumb and body, dizziness, cervical myalgia/paresthesia, numbness and burning of the left shoulder, and numbness of the arm and tightness of the scapula.

It is important to note that comparison of events between groups is confounded by the fact that the categories reported are not consistent between the groups as is noted in the table above. However, most of these events are consistent with involvement of cervical disc disease. It is not clear whether these events were new or exacerbations of existing disease in either group and to what extent if any these events resolved or persisted.

Spinal Adverse Events

A total of 18 spinal events occurred in 17 patients in the investigational group (6.2%).

The following table summarizes the spinal adverse events that occurred during the study. Following the table is a detailed description of each of the spinal adverse events.

Event Type	Investigational	Control
Lumbar Spine	14	15
Thoracic Spine	1	4
Cervical Spine	3	13

Investigational

The most frequently reported events were lumbar-associated (14 events). These events included the following: three herniated discs, three reports of degenerative disc disease, two cases of spondylolisthesis/listhesis, two cases of stenosis, two disc bulges, one collapsed disc, one transition syndrome, and one post-laminectomy syndrome. Additionally, there was one disc herniation in the thoracic spine.

Also reported were three cervical-associated events. There was one report of disc herniation and one report of degenerative disc disease that occurred adjacent to the level at which the investigational device was implanted. One patient was reported to have a mild, degenerative spondylolisthesis at the treated level 12 months postoperatively and was continuing treatment with medications and physical therapy.

Control

In the control group, there were 32 spinal events noted in 30 patients (11.3%). Again, the most frequently reported events were lumbar-associated (15 events). These included seven reports of degenerative disc disease with stenosis or herniated disc, five herniated discs, two cases of stenosis, and one spondylosis. There were four thoracic-associated events: three herniated discs and one degenerative disc disease.

In addition, there were 13 cervical-associated events reported. These included three herniated discs, three herniated discs and disc bulges, and two cases of stenosis. Finally, each of the following occurred once: degenerative disc disease, ossification at an adjacent level, stenosis/herniated disc, kyphosis/stenosis, and spurs/disc bulge.

Trauma

A total of 69 trauma events occurred in 59 patients in the investigational group (21.4%). In the control group, there were 47 trauma events noted in 40 patients (15.1%).

The following table summarizes the trauma events that occurred during the study. Following the table is a detailed description of each of the trauma events.

	Investigational	Control
Motor Vehicle Accidents	22	10
Falls	13	18
Work Related	5	7
Assaults	2	0
Lacerations	3	0
“Jarred” Necks	2	0
Back Injuries	2	0
Sudden Movement	2	0
Over-activity	0	2
Meniscal Tears	0	2
Other*	16	8

*Events that occurred in only one patient.

Investigational

The most frequently reported events were motor vehicle accidents (22 events) and falls (13 events). Also reported were five work-related injuries, two assaults, three lacerations, two patients with “jarred necks”, two reports of lifting injuries, two back injuries, and two events

caused by sudden movements. Additionally, there were 16 events that occurred only once and were reported as trauma. These included: dog bite; snake bite; running injury; right hand tendon injury; torn left wrist tendon; rotator cuff tear; tractor accident; amusement park ride trauma; injury when hit with a heavy gate; “head popped”; severed digits; contact injury (hugged too hard); an injury due to yoga; and injuries to the eye, hand, and shoulder.

Control

Similarly, the most frequently reported were falls (18 events) and motor vehicle accidents (10 events). There were seven reported work-related injuries. Two injuries were secondary to over activity and exercise-induced, and there were two meniscal (knee) tears. Additionally, there were eight events that occurred only once and were reported as trauma. These included: eye injury, hernia, muscle strain, neck strain, pulled muscle, “jammed head”, paddle boat falling on the shoulder, and a finger slammed in the door.

Neck and/or Arm Pain

A total of 190 events occurred in 138 patients in the investigational device group (50.0%). By comparison, a total of 173 neck and/or arm pain events occurred in 127 patients in the control group (47.9%).

The following table summarizes the neck and/or arm pain events that occurred during the study. Following the table is a detailed description of each of the neck and/or arm pain events.

Event Types	Investigational	Control
Neck Pain	42	58
Shoulder Pain	35	24
Neck/Shoulder Pain	20	11
Arm Pain	11	20
Neck Spasms	10	9
Neck/Arm Pain	10	10
Interscapular/Scapular Pain	8	9
Shoulder/Arm Pain	7	0
Epicondylitis	5	3
Neck/Headache	4	4
Trapezius	4	8
Wrist Pain	3	0
Rotator Cuff	3	6
Neck/Shoulder/Arm	3	2
Hand Pain	3	0
Shoulder Tendonitis	2	0
Elbow Pain	0	2
Shoulder Impingement	0	2
Other*	19	5

*Events that occurred in only one patient

Investigational

The events included the following: 42 neck pain; 35 shoulder pain; 20 neck and shoulder pain; 11 arm pain; 10 neck spasms; 10 neck and arm pain; eight interscapular/scapular pain; seven shoulder and arm pain; five epicondylitis; four neck and headache; four trapezius pain; three wrist pain; three rotator cuff events; three neck, shoulder and arm pain; three hand pain; and two shoulder tendonitis.

Additionally, there were 19 events that occurred only once and were reported as neck and arm pain. These included acromial/clavicle pain; deltoid pain; elbow pain; jaw pain, muscle strain; cervical strain; elbow, wrist and hand pain; neck and ear pain; scapular pain with migraine headache; neck, shoulder, and upper back pain; shoulder, arm and hand pain; submastoid and arm pain; T1-T2 pain; shoulder joint degeneration; neck and shoulder pain with arm numbness; neck and scapular pain; neck and thoracic pain; shoulder impingement; and trapezius pain.

Control

These events included 58 neck pain; 24 shoulder pain; 20 arm pain; 11 neck and shoulder pain; 10 neck and arm pain, nine neck spasms; nine interscapular/scapular pain; eight trapezius pain; six rotator cuff events; four neck pain with headache; three epicondylitis; two neck, arm, and shoulder pain; two elbow pain; and two shoulder impingement.

Additionally, there were five events that occurred only once and were reported as neck and arm pain. These included neck and scapular pain, radiating pain, neck and upper back pain, glenohumeral joint pain, and clavicle/scapular pain.

Other Pain

A total of 88 events classified as “other pain” occurred in 69 patients in the investigational group (25.0%). A total of 68 events classified as “other pain” occurred in 56 patients in the control group (21.1%).

The following table summarizes the other pain events that occurred during the study. Following the table is a detailed description of each of the other pain events.

Event Types	Investigational	Control
Back Pain	28	27
Headache	22	16
Back/Leg Pain	10	4
Hip Pain	7	2
Knee Pain	6	4
Leg Pain	4	3
Thoracic Pain	4	0
Back/Hip Pain	2	3
Other*	5	9

*Events that occurred in only one patient.

Investigational

The most frequently occurring categories were back pain (28 events) and headaches (22 events). In addition, there were 10 reports of back and leg pain, seven reports of hip pain, six reports of knee pain, four reports of leg pain, four reports of thoracic pain, and two reports of back and hip pain. Each of the following types of pain was reported once: back/thoracic, hip/leg, bursitis, knee/ankle, and flank.

Control

Again, the most frequently reported events were back pain (27 events) and headaches (16 events). In addition, there were four reports of back and leg pain, four reports of knee pain, three reports of back and hip pain, three reports of leg pain, and two reports of hip pain. Each of the following types of pain was reported once: flank/pelvis, flank/knee, foot, abdominal, back/pelvis, and leg/hip. Finally, there was one occurrence each of thoracic pain, incision pain, and sacroilitis.

None of the events in this category were related to the cervical spine.

Cancer

The incidence rate of cancer in the investigational group was 1.8% (5 patients). This compares to a 0.8% rate (2 patients) for the control group. The rates were not statistically different.

The following table summarizes the cancer events that occurred during the study. Following the table is a detailed description of each of the cancer events.

	Investigational	Control
Cancer	5	2

Investigational

Of the five investigational patients, the first was a non-Hodgkin's lymphoma, which occurred approximately 26 months following surgery. The diagnosis was made based on a biopsy of a nasal mass found on a CT performed to investigate the patient's complaint of hearing loss. The second report of cancer was basal cell carcinoma reported at the patient's 24-month visit. The third report of cancer occurred at 17 months postoperative when the patient was diagnosed with colon cancer. The fourth report of cancer was reported at the 24-month visit, when the patient reported being seen by an ENT who diagnosed a thyroid cancer. The fifth cancer reported in the investigational group was a breast carcinoma occurring approximately 17 months following surgery.

Control

Of the two patients in the control group for whom a cancer was reported, the first was at approximately 23 months following surgery when the patient reported a reoccurrence of skin cancer. The second report of cancer in the control group occurred about 7 months following surgery when the patient was found to have a brain tumor, i.e. a low grade, non-metastatic astrocytoma unrelated to the cervical spine.

FDA Question for Panel:

There were five incidences of cancer in the treatment group as opposed to two incidences in the control group. Two of the patients developed cancer in the first 12 months of follow-up and the remaining three patients during the 12 to 24 month follow-up. Considering the concerns with metal on metal devices (e.g., particulate wear generation, particulate migration, etc.) and the metal ion testing, please discuss whether this raises safety concerns with the investigational device. Also, should there be a section in the labeling discussing this issue?

Deaths

There were no deaths in the investigational device group. In the control group, there were three deaths reported (1.1%).

The following table summarizes the deaths that occurred during the study. Following the table is a detailed description of each death.

	Investigational	Control
Death	0	3

Control

The first was a 68-year-old female who suffered a fatal myocardial infarction at home 3 months following surgery. The patient had a previous history of cardiac problems. The second death

occurred approximately 11 months postoperatively when the patient suffered an acute myocardial infarction. The patient was taken to the emergency room, where he received multiple treatments for cardiac arrest. The patient subsequently died as a result of the cardiac arrest. The third death occurred at about 24 months postoperative when the patient went to the emergency room complaining of right arm pain, chest pain, and shortness of breath. The patient was kept overnight and discharged the next day. The patient was instructed to take an aspirin a day. A few days later, the patient had a fatal cardiac arrest.

Radiologist Findings of Damaged Devices

In the investigational group, the radiographic reviewers did not note any findings of implant bending, breakage, migration, or fracture. In the control group there were eight reports of damaged devices.

The following table summarizes the radiologists finding of damaged devices. Following the table is a detailed description of each damaged device finding.

	Investigational	Control
Damaged Devices		
Implant Migration	0	3
Broken/Fracture Bone Graft	0	5

Control

In the control group, implant migration was noted in three patients [redacted]. None of [redacted] resulted in an apparent second surgical procedure as Patient [redacted] did have a procedure classified as “other” to evacuate a hematoma following the study s [redacted].

In the control group, five patients [redacted] were reported to have broken or fractured bone grafts. T [redacted] plants removed or replaced in subsequent revision (one) and removal (two) procedures. The revision procedure was due to the need to fuse an adjacent level. One of the removals was elective, while the other removal occurred in a subsequent procedure related to an infection.

Summary

With the possible exception of the incidence of cancer in the investigational group, the types and rates of adverse events appear to be similar in the two groups.

FDA Question for Panel:

Please discuss whether the clinical data in the PMA provide reasonable assurance that the proposed device is safe. If not, what additional data are needed?