Proposed Decision Memo for Lumbar Artificial Disc Replacement (LADR) (CAG-00292R)

Decision Summary

The Centers for Medicare and Medicaid Services (CMS) proposes that LADR is not reasonable and necessary for the Medicare population over sixty years of age. Therefore, CMS is proposing that Section 150.1 of the Medicare National Coverage Determination (NCD) Manual be amended to reflect the proposed change from non-coverage for a specific LADR implant to non-coverage for the LADR procedure for the Medicare population over sixty years of age. For Medicare beneficiaries sixty years of age and under, there is no national coverage determination, leaving such determinations to be made on a local basis.

We are requesting public comments on this proposed determination pursuant to §1862(1) of the Social Security Act. We are particularly interested in comments that include evidence we did not review or that assess how we evaluated the evidence included. After considering the public comments and any additional evidence we will make a final determination and issue a final decision memorandum.

Back to Top

Proposed Decision Memo

TO: Administrative File: (CAG-#00292R)
Lumbar Artificial Disc Replacement

FROM:

Steve Phurrough, MD, MPA
Director
Coverage and Analysis Group

Marcel E. Salive, MD, MPH
Director
Division of Medical and Surgical Services

Deirdre O'Connor
Lead Health Policy Analyst, Division of Medical and Surgical Services

Jyme Schafer, MD, MPH
Lead Medical Officer, Division of Medical and Surgical Services

SUBJECT: Proposed Coverage Decision Memorandum for Lumbar Artificial Disc Replacement (LADR)
DATE: May 25, 2007

I. Proposed Decision

The Centers for Medicare and Medicaid Services (CMS) proposes that LADR is not reasonable and necessary for the Medicare population over sixty years of age. Therefore, CMS is proposing that Section 150.1 of the Medicare National Coverage Determination (NCD) Manual be amended to reflect the proposed change from non-coverage for a specific LADR implant to non-coverage for the LADR procedure for the Medicare population over sixty years of age. For Medicare beneficiaries sixty years of age and under, there is no national coverage determination, leaving such determinations to be made on a local basis.

We are requesting public comments on this proposed determination pursuant to §1862(1) of the Social Security Act. We are particularly interested in comments that include evidence we did not review or that assess how we evaluated the evidence included. After considering the public comments and any additional evidence we will make a final determination and issue a final decision memorandum.

II. Background

Millions of Americans suffer from pain-related problems (Salovey, Seiber et al. 1992). Low back pain is a common condition, with sixty to eighty percent of U.S. adults afflicted at some time during their life (U.S. Preventive Services Task Force 1996). Low back pain can be defined as symptoms of pain, muscle tension, or stiffness localized below the costal margin and above the inferior gluteal folds, with or without leg pain (Manek, MacGregor 2005). Low back pain can be thought of as being either nonspecific or specific. In specific types of low back pain, the symptoms are caused by pathological conditions such as spinal fractures, cancer, or infection and can be identified and treated appropriately (Manek, MacGregor 2005). Approximately 90% of low back pain is of the nonspecific type (Manek, MacGregor 2005). In nonspecific low back pain, most patients' symptoms resolve satisfactorily within a relatively short time span. In the 5 – 10% of patients whose pain does not satisfactorily resolve, the symptoms can be disabling. Some psychosocial risk factors for the progression to chronicity have been identified (Manek, MacGregor 2005). In general, the social and economic impact of chronic pain is enormous (Salovey, Seiber et al. 1992).

Discovering the cause for nonspecific low back symptoms remains challenging. Haldeman stated "…we do not know the origin of low back pain in the majority of cases…" and attributes this conundrum to the unique anatomic complexity of the spine (Haldeman 1999). Neurophysiologic mechanisms of pain sensation are poorly understood, adding to the difficulty in localizing the pain source (Haldeman 1999). Frequently, persistent low back pain is attributed to a damaged intervertebral disc, which bears some of the highest loads in the human body and is almost avascular (Huang, Sandhu 2004). Disc damage, or degeneration, can occur as an ongoing process where ultimately the disc's reparative capacity is overwhelmed, leading to continued changes. Huang and Sandhu stated, "it is not surprising that DDD [degenerative disc disease] is a common phenomenon in middle age and a universal condition in old age." While from a simple mechanical aspect it could be hypothesized that DDD is a cause for pain, disc degeneration is also observed in individuals without pain (Boden, David et al. 1990).

Initial treatment of pain believed to be caused from degenerative disc disease is conservative care. Conservative care can include physical therapy, manipulation, massage, pain medications, and exercise. The majority of patients will have acceptable results with a non-surgical approach. When patients fail conservative care, surgery becomes an option. Until recently in the United States, surgical options available for degenerative disc disease have ranged from discectomies (open or microsurgical) to percutaneous nucleotomies, chemical and thermal nucleolysis and/or spinal fusion (Gibson, Wassell 2005). Spinal fusion has been the predominant surgical treatment for degenerative disc disease (DDD) that does not respond to other treatments. Fusion proposes to relieve pain by eliminating motion in the area of the disc space and/or by disc mechanical load reduction. Nevertheless, the indications for lumbar spinal fusion are variable and not clearly defined (Krismer 2002). These different opinions concerning the indications for back surgery are reflected in the significant regional variation of rates of surgery, surgical techniques used, technical success and rate of fusion (Gibson, Wassell 2005). Satisfactory clinical outcomes can range from 16 to 95% (Gibson, Wassell 2005). Short term relief of pain may perhaps occur with the various types of fusion procedures, but long-term results remain controversial (Bertagnoli, Kumar 2002). Suspected problems include accelerated degeneration of the adjacent lumbar segments, pseudoarthrosis, spinal stenosis and persistent or recurrent low-back pain. In an attempt to overcome these potential long-term problems, the idea of a total artificial disc replacement as a treatment for pain believed secondary to degenerative disc disease has been proposed as an alternative to spinal fusion. As possible added benefits, it has been postulated that total disc replacement may have a protective role on the facet joints, and restore lumbar segment motion (Bertagnoli, Kumar 2002). The artificial disc concept is not new. In the late 1960's, Fernstrom explored the possibility of replacing the intervertebral disc with an artificial disc. Much research and development work has been done since then. Of the two lumbar artificial discs that are currently FDA approved, the Charite disc is the third modification of a device first developed in 1982 by Buttener-Janz and Schellnack at the Charite Clinic in the former East Germany and the ProDisc®-L disc is the second generation of the device designed in the late 1980's by Marnay.

Intervertebral disc replacement design has been problematic due to the three-column structure of the spine, and the three separate joints at each level. The disc is not a true joint, and functions in both mobility and damping, with the center of rotation moving constantly along three axes (Gunzburg, Mayer et al. 2002). Huang and Sandhu suggest the ideal disc replacement would perform the functions of the replaced native disc, which include preservation of physiologic range of motion, transmission of compressive loads across the disc space, protection of the posterior elements (facets) from abnormal loads, and then to function for many years. In general, the current replacement discs that are either approved or under FDA approved trials in the US have metal endplates that affix to the vertebral bony endplates with some mechanism between these two plates that allows for motion in various planes. The ProDisc®-L and the Charite have similar modular designs but differ in the mechanical design mainly in how the metal endplates affix to the vertebral body and the fixation of the poly inlay. The ProDisc®-L disc has two metal endplates with ultimately fixation to the vertebral body through bony ingrowth and initial stabilization provided by a centrally located keel. The ultra-high molecular weight polyethylene inlay locks in place to the inferior endplate thereby producing a semi-constrained device. The other disc implants in development in the United States are somewhat similar but can vary in material (metal on polymer or metal on metal), motion design, and method of fixation to vertebral endplate (Santos, Polly et al. 2004). In 2004 Anderson and Rouleau offered, "The current designs are diverse and, thus far, the effects of their individual characteristics on results are unknown." The Food and Drug Administration summary noted "The ProDisc®-L total disc replacement has been commercially available in markets outside of the United States since 1990" (FDA Summary of Safety and Effectiveness Data for Expedited Premarket Approval (PMA) 2006).

The surgical procedure for disc replacement involves an anterior approach for exposure of the spine. With this approach, complications of vessel injury can occur and have the potential to be life threatening (Santos, Polly et al. 2004). On revision surgery, Santos et al. stated, “Revision surgery for a failed disc arthroplasty is life threatening. Dealing with the scarring around the great vessels is the main challenge. Indeed, the location of vital vascular structures may make it altogether impossible to perform such anterior abdominal exposures.” Other postoperative difficulties such as infection, persistent pain, instability, and osteolysis can occur (Santos, Polly et al. 2004).

III. History of Medicare Coverage

On May 16, 2006, CMS issued a NCD (CMS NCD Manual Section 150.10) for LADR. The coverage decision was focused on the Charite™ lumbar artificial disc because it was the only lumbar artificial disc with FDA approval at that time. The current policy for LADR replacement is as follows:

LADR with the Charite™ lumbar artificial disc is not reasonable and necessary for the Medicare population over 60 years of age; therefore, LADR with the Charite™ lumbar artificial disc is non-covered for Medicare beneficiaries over 60 years of age. For Medicare beneficiaries 60 years of age and younger, there is no national coverage determination, leaving such determinations to continue to be made by the local contractors. Medicare coverage under the investigational device exemption (IDE) for other lumbar artificial discs in eligible clinical trials is not impacted.

In the decision memorandum for LADR issued on May 16, 2006, CMS stated, “CMS is aware that there are several other disc technologies in FDA investigational device exemption clinical trials in the United States. As previously stated, CMS is evaluating LADR with a focus on the Charite lumbar artificial disc in this analysis, since this was the only disc implant that had FDA approval at this time. However, we anticipate that when other lumbar spinal disc implants receive approval from the FDA that CMS will, by external request or internal direction, open this NCD for reconsideration with a thorough review of the evidence for each new disc implant.”

Benefit Category

Medicare is a defined benefit program. An item or service must fall within a benefit category as a prerequisite to Medicare coverage. §1812 (Scope of Part A); §1832 (Scope of Part B); §1861(s) (Definitions of Medical and Other Health Services). LADR would be eligible for coverage under Part B, as physician services, under §1861(s)(1) and (2)(A) and under Part A, inpatient hospital services, under §1861(b). This may not be an exhaustive list of all applicable Medicare benefit categories for this item or service.

IV. Timeline of Recent Activities

November 28, 2006	CMS initiates opening reconsideration of NCD for LADR. Initial 30-day public comment period begins.
December 28, 2006	Initial 30-day public comment period closes.
January 8, 2007	Meeting with Synthes

V. Food and Drug Administration (FDA) Status

The FDA approved the PRODISC®-L Total Disc Replacement on August 14, 2006 (http://www.fda.gov/cdrh/pdf5/p050010a.pdf).

The FDA approval letter stated, “This device is indicated for spinal arthroplasty in skeletally mature patients with degenerative disc disease (DDD) at one level from L3-S1. DDD is defined as discogenic back pain with degeneration of the disc confirmed by patient history and radiographic studies. These DDD patients should have no more than Grade 1 spondylolisthesis at the involved level. Patients receiving the PRODISC®-L Total Disc Replacement should have failed at least six months of conservative treatment prior to implantation of the PRODISC®-L Total Disc Replacement.”(FDA Approval Letter, August 14, 2006)

VI. General Methodological Principles

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. The critical appraisal of the evidence enables us to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for patients. An improved health outcome is one of several considerations in determining whether an item or service is reasonable and necessary.

A detailed account of the methodological principles of study design that the agency utilizes to assess the relevant literature on a therapeutic or diagnostic item or service for specific conditions can be found in Appendix A. In general, features of clinical studies that improve quality and decrease bias include the selection of a clinically relevant cohort, the consistent use of a single good reference standard, and the blinding of readers of the index test, and reference test results.

Public comment sometimes cites the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as the results of individual practitioners or patients are less rigorous and therefore less useful for making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

VII. Evidence

A. Introduction

A summary of the evidence used to arrive at the proposed determination is provided. This summary represents the evidence relating to the treatment of pain from degenerative disc disease with LADR with the ProDisc lumbar artificial disc and includes a clinical trial, case series reports, and technical reviews. The evidence CMS examines has as its focus health outcomes, or, the benefits and harms of a particular treatment. Outcomes that are usually heavily weighted by CMS - morbidity and mortality - are difficult to examine in the context of treatment for chronic low back pain which is a symptom, not a disease. In chronic low back pain, sustained improvement in pain perception and a reduction in the pain-related functional restriction are generally the focus of study outcomes. Measuring a reliable improvement in chronic pain is problematic as pain is subjective and is particularly responsive to the placebo effect; therefore, clinical trials with appropriate controls utilizing independently assessed validated instruments are most heavily weighted. The measurement of treatment effect for low back pain has shifted from physician-based assessment (with outcomes of excellent, good, fair, and poor) to a patient-based self-report of pain and disability (Hagg, Fritzell et al. 2003).

Treatment effect in chronic low back pain is measured with patient-based, multi-item instruments. Two instruments validated for measurement of back pain are commonly used in the assessment of low back pain from degenerative disc disease (Hagg, Fritzell et al. 2003).

The Oswestry Disability Index (ODI) is a condition-specific outcome measure used in the management of spinal disorders. The measure is an indication of the extent to which a person's functional level is restricted by pain. The other commonly used measure in chronic back pain treatment effect is the visual analogue scale (VAS), which is a method to assess pain intensity. With the use of these instruments for measurement, a consideration must be given to the clinical meaning of a change in the score (or, for a change in instrument score to be clinically meaningful the patient should experience a change in how he feels or functions). Other considerations include the error of measurement of the instrument used and the clinical importance of a statistically significant score change. In a 2003 study by Hagg of 289 patients treated surgically or non-surgically in a randomized controlled trial, the standard error of measurement of the ODI was 4 units, with a 95% tolerance interval of 10, and the minimum difference that appeared clinically important was 10 units (Hagg, Fritzell et al. 2003). The minimal clinically important difference of VAS back pain was 18 – 19 units with a 95% tolerance interval of 15. It was interesting to note that in this study, improvement after treatment tended to occur to a greater extent in sleep disturbance, ability to do usual things and psychological irritability, but to a lesser extent in the ability to sit, stand and lift.

Some investigators have used the Stauffer Coventry classification, or some modification thereof to measure results. The criteria for clinical results for the Stauffer and Coventry classification are provided in Table 1 (Sott, Harrison 2000).

Table 1 – Stauffer Coventry Classification

Pain relief (%)	Return to work	Physical restriction	Use of analgesics
Good 76 – 100	Yes	No or slight	No
Fair 26 – 75	Yes, with limitations	Yes, limited activities	Frequent (mild)
Poor < 25	No, disabled	Yes, greatly limited	Regular (strong)

Additionally, other quality of life measures are sometimes used. The SF-36 Health Survey, a 36 question form that measures general health status, can be used. Of the 8 health profiles that are included in this survey, only one or two components may be reported, such as the physical functioning composite score or the mental health composite score.

Some studies have reported range of motion as an outcome. Physiologic segmental mobility, as measured by range of motion, is viewed by some artificial disc proponents as an important design feature of the disc. This view is based on the premise that fusion surgery alters normal motion of the adjacent level disc, resulting in an increased likelihood of disease in those adjacent discs. Conversely, these proponents postulate that motion preservation by the artificial disc will prevent this. In our review of the literature, we were unable to find evidence that the theoretical mobility provided by the artificial disc directly correlates to a benefit in how the patient feels or functions, making the clinical significance of post treatment range of motion unclear. In addition, we are unable to identify any clinical evidence that supports the premise that segmental mobility prevents adjacent level disease. Therefore, CMS does not consider post treatment range of motion an important clinical outcome of interest in this memorandum.

Well-designed clinical trials can provide the strongest evidence for treatment effect. Clinical trials can be designed to show superiority, a priori, where the superior clinical performance of the investigational agent as compared to the control agent is anticipated. When the investigational agent is believed to have comparable efficacy to the control, but has other advantages, for example fewer adverse events or less cost, a noninferiority trial is an option. In a noninferiority trial, the aim is to demonstrate that the investigational agent is not worse than the control by a certain pre-specified margin, referred to as the delta. In the statistical approach for noninferiority analysis, the delta is compared with the one-sided 95% confidence interval for the difference between the success rate point estimates of the investigational agent and control. If the lower bound of this one-sided confidence interval is less than the delta, then the statistical definition of noninferiority is met.

B. Discussion of evidence

1. Question:

The development of an assessment in support of Medicare coverage decisions is based on the same general question for almost all requests: "Is the evidence sufficient to conclude that the application of the item or service under study will improve health outcomes for Medicare patients?" For this NCD, the question of interest is:

Is the evidence sufficient to conclude that LADR with the ProDisc lumbar artificial disc will improve health outcomes in the Medicare population with low back pain due to degenerative disc disease?

2. External technology assessment

CMS did not commission an external technology assessment (TA); however, an external TA was identified on the topic of Artificial Vertebral Disc Replacement.

In April of 2005, the Blue Cross Blue Shield Technology Evaluation Center (TEC) published a TA titled, Artificial Vertebral Disc Replacement. Artificial Vertebral Disc Replacement met only one of five of the TEC criteria. The TEC determined "…the use of artificial vertebral discs for degenerative disc disease does not meet the TEC criteria." The following criteria were not met: 1) The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes; 2) The technology must improve the net health outcome; 3) The technology must be as beneficial as any established alternatives; and, 4) The improvement must be attainable outside the investigational setting.

In summary the TEC report stated the following.

"…the evidence supporting the effectiveness of the Charite artificial disc is limited. Case series provides little evidence of efficacy, particularly in the case of back pain due to degenerative disc disease, where outcomes can be influenced by patient selection, placebo effects, or natural history."

It was further stated:

"The only randomized, controlled trial has several methodologic issues that make it difficult to interpret results." "…A noninferiority trial design implies that there is a trade-off between efficacy outcomes and some other advantage of a new technology, for example, morbidity or invasiveness, such that a less-stringent threshold for efficacy is acceptable. However, at this time, no such advantage has been demonstrated for the Charite artificial disc."

Another concern identified was "…that the lack of a prespecified analysis plan, unexplained closure of the database before all patients reached completion, and lack of intent-to-treat analysis may cast some doubt on the analysis."

The TEC conclusion was, "Given the broader clinical context, and the concerns with the sole randomized, controlled trial, the evidence is not sufficient to conclude that the use of artificial vertebral disc improves health outcomes." They further concluded that "The evidence is insufficient to determine whether the use of artificial vertebral discs improves net health outcome or whether they are as beneficial as any established alternative."; and expressed concern "Whether the use of artificial vertebral discs improves health outcomes has not been established in the investigational settings."

The Medical Advisory Panel for the BCBS TEC reaffirmed the decision that artificial lumbar disc for DDD does not meet TEC criteria in March of 2007 (BCBS TEC MAP March 2007).

3. Internal technology assessment

The evidence summary and analysis in the original decision memorandum on LADR (available at http://www.cms.hhs.gov/mcd/viewdecisionmemo.asp?id=170) is incorporated into this document by reference. CMS performed an additional literature search utilizing PubMed for randomized (and nonrandomized) controlled trials (RCTs), cohort or case-control studies, case series studies and systemic reviews evaluating the use of ProDisc lumbar artificial disc replacements for the treatment of degenerative disc disease. The literature search was limited to the English language and specific to the human population, but included studies conducted in all countries, including the United States (see evidence tables in Appendix B). Public access information from the FDA website was also used.

Evidence for the ProDisc lumbar artificial disc came from the FDA PMA Application clinical trial and several case series reports.

Evidence Summary

Observational Studies

A 2002 abstract by Thierry Marnay, ProDisc inventor, reported certain 7 – 11 year results with Prodisc (Marnay 2002). From 1990 to 1993, ninety-three prostheses were implanted in 64 patients from L2 to S1, of which thirty-nine had one level, twenty-one had 2 levels and four had 3 levels implanted. The age range of patients was not provided. Patients had failed conservative care for chronic back pain. It was stated, “The last follow up patients were at 8.6 years average of post surgery.” The time schedule for follow up evaluations was not provided. The VAS average preoperative score was 8.5 and then 3.0 at the final follow up (range of scores or point of time measurement not listed). At the long-term follow up, 65% of the patients reported that they were “entirely satisfied”, 28% “satisfied” and only 7% were “not satisfied”. The overall ODI score average was 8.3 (preoperative score, follow-up time not listed). During follow-up, five patients had fusion due to ongoing pain, one had a vascular complication and 2 had temporary sexual dysfunction. The author concluded “…that the Prodisc can remain mechanically stable and provide significant pain relief while maintaining motion in patients at 7 to 10 years follow-up.” It is not clear how many patients were lost to follow-up.

In 2002 Mayer reported on the surgical technique for total lumbar disc replacement and included preliminary results for 34 patients (Mayer, Wiechert et al. 2002). Average age was 44 (range 25 to 65 years). The main indication was degenerative disc disease. Exclusions included spondylolisthesis, spinal stenosis, significant osteoarthritis of the facet joints, deformities, infection or tumor, unwillingness to comply with study requirements regarding follow-up visits and radiological controls, previous fusion attempts in the affected levels, pregnancy and incomplete workers' compensation procedures. Twenty-six (76.5% of the original 34) patients had at least one visit for evaluation. Results revealed a mean operative time of 130.9 minutes (range 88 to 300minutes) and average blood loss was 117 ml per level (range 30 – 350 ml). The mean VAS preoperative score was 6.3 and was reduced at the 12 month visit to 3.4 (number of patients not given, standard deviation not given). The ODI ranged from 1 to 32 points before surgery (average 19.1, standard deviation 7.4 points). The average score at 12 months was 7.2 ("The change in postoperative score ranged from 27 points reduction to an increase of 12 points (standard deviation 9.6 points)"). There was no difference noted in results between one and two level implantation. Three complications related to the surgical procedure were reported.

Tropiano reported a prospective analysis following 53 patients implanted by a single surgeon for a minimum of one year (range 1 to 2 years) (Tropiano, Huang et al. 2003). Patients had a mean age of 45 years (range 28 – 67 years) and included 18 men and 35 women. Diagnoses included degenerated disc and failed back surgery. Patients were excluded if they had facet degeneration, a history of abdominal or retroperitoneal surgery near planned surgical approach, osteoporosis or osteopenia, structural spinal deformities, or an absence of posterior elements due to previous surgery. Eleven patients had 2 discs implanted and two patients had three levels implanted. Mean operative time was 104 minutes (range 32 – 250 minutes), with a mean hospital stay of 9 days (range 4 – 31 days). Patients received low molecular weight heparin as prophylaxis for 21 days postoperatively. Patients began physiotherapy one month after surgery and also advanced to unrestricted activities as tolerated at that time. Clinical outcomes included back and leg pain as measured by the VAS (modified version); pain intensity measured on a 10 point scale from severe to none; ODI; quality of life measured as normal, slightly limited, hindered, or severely limited/impossible; return to work measured as normal, slightly limited, hindered, or severely limited/impossible; and patient satisfaction measured as entirely satisfied, satisfied or not satisfied. Results revealed 100% of patients were entirely satisfied or satisfied; 72% of patients resumed work and activities of daily living with 28% being slightly limited (though 7 of these patients on workers' compensation said they could not work); patients improved significantly in VAS lumbar and radicular pain (to a mean score at 1.4 year follow-up of 1.3 +/- 1.78 for VAS lumbar and 1.9 +/- 2.59 VAS radicular); ODI improved from a preoperative mean score of 56 to a mean score of 14 at 1.4 years. Radiographic results revealed flexion-extension of 8 degrees for those implanted at L5-S1, and 10 degrees for those implanted at L4-L5. Clinical results of the single and multilevel replacements were equivalent. Complications occurred in 9% of patients and included vertebral body fracture, radicular pain, implant malposition, and retrograde ejaculation. Complications necessitated reoperation in three patients.

In 2005 Tropiano reported on clinical and radiographic results in 55 ProDisc patients (64 patients initially in the study), with a mean duration of follow-up of 8.7 years (range 7 to 11 years) (Tropiano, Huang et al. 2005). Patients had a minimum of 6 months nonoperative treatment prior to procedure. Exclusion included facet arthrosis, central or lateral recess stenosis, osteoporosis, sagittal or coronal plane deformity, and absence of posterior elements. Average patients age was 46 (range 25 to 65). A report of clinical results included, "Clinical results were evaluated by assessing preoperative and postoperative lumbar pain, radiculopathy, disability, and modified Stauffer-Coventry scores." Stauffer–Coventry score (0-20 points) increased from an average of 7.04 pre-op to 16.1 post-op. Low back pain, lower limb pain, and impairment (all measured on 3 point scales) decreased at post-operative measurement. The authors stated, "Thirty-three of the fifty-five patients with sufficient follow-up had an excellent result, eight had a good result, and fourteen had a poor result." Seven patients had both disc replacement and an adjacent fusion during the same operation. Five patients had approach related complications.

Bertagnoli reported on a case series of 118 patients age 18 to 60 with low back pain with or without radicular symptoms resulting from degenerative disc disease from L3 to S1, by a single surgeon using the ProDisc (Bertagnoli, Yue et al. 2005a). Patient's pre-op assessment included plain radiographs, MRI and CT scans. Discography was used in selected circumstances. Exclusion criteria included patients with spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, pregnancy, facet arthrosis, inadequate vertebral endplate size, more than one level of spondylosis, neuromuscular disease, workers' compensation, spinal litigation, body mass index greater than 35, and any isthmic or degenerative spondylolisthesis greater than Grade 1. Patients had failed conservative treatment for a minimum of 9 months. The ODI, VAS, and measures of back pain and radicular pain (unclear if outcomes were validated) were chosen as outcomes. To assess changes over time, statistical methods used a longitudinal approach with general linear models (GLM) for continuous variables and generalized estimating equations (GEE) for patient satisfaction and back pain. Of the original 118 patients only 104 patients with complete data were analyzed. This interesting analysis showed that at 24 months, 91% had either occasional pain or no pain, with significant decreases in Oswestry disability scores that were sustained at 24 months (53% to 29%) and VAS change (7.6 to 3) was also sustained at 24 months. Medication usage revealed that 83% pre-operatively did not use narcotics and 90% post-op did not use narcotics. There were no device-related complications but several approach-related complications (3 hematomas and 1 retrograde ejaculation) that resolved, one complication of persistent leg pain that required posterior exploration and decompression, which revealed posterior subarticular stenosis.

In a 2005 case series Bertagnoli reported on patients with multilevel ProDisc implants (Bertagnoli, Yue et al. 2005b). This analysis included 25 patients (15 male, 10 female) with a minimum follow-up of 2 years, implanted by a single surgeon at a single site. Patients age 18 to 60 (median age 51) with disabling low back pain and minimal radicular pain with multiple lumbar spondylosis from L1 to S1 (confirmed by MRI and discogram/CT) were included in this study. Patients had failed 9 months of conservative treatment prior to implantation. Exclusion criteria were similar to other Bertagnoli reports, therefore did not include those with significant facet arthropathy, workers' compensation, or spine litigation. Fifteen patients had 2 level implants and 10 patients had 3 level implants. VAS, ODI, leg and back pain percentages and radiographic data was collected at 3,6,12, and 24 months. The average operative time for a 2 level surgery was 135 minutes and for three levels was 184 minutes. Blood loss for a two level surgery averaged 275 ml and 350 ml for a three level surgery. Patients were discharged approximately 3.5 days post-op. For statistical analysis, repeated measures general linear models (GLM) were used for continuous variables and generalized estimating equations were used for dichotomous variables. Oswestry scores decreased from 65% to 21% at 24 months ($p < 0.001$) and VAS scores decreased from 8.3 to 2.1. Before surgery, all patients reported back pain. At the 24 month follow-up 92% of patients reported no or occasional back pain. Forty-eight percent of all patients had no or occasional leg pain pre-operatively, increasing to 100% at 24 months follow-up. The rate of patient satisfaction was 92% at 2 years. Post-op radiographs were obtained, with no correlation between clinical outcomes and pelvic incidence, tilt or sacral slope. Complications included a case of subsidence in a 36 year old male with no prior history of osteoporosis and a case of anterior extrusion of the polyethylene core in a patient who fell off a bicycle. There were no cases of vascular injury or neurologic injury. The authors concluded, "We think that these excellent results are a direct result not only due to the qualities of the implant, but moreover, of careful patient selection by an experienced low back surgeon."

In 2006 Bertagnoli reported on 20 patients, age 18 to 67 (median age 50, number of patients over 60 not reported), who were treated for symptomatic adjacent-segment degeneration after remote lumbar fusion (Bertagnoli, Yue et al. 2006a). Studies were performed by a surgeon at a single center. Pre-op studies included MRI, CT, and discography. Exclusion criteria were circumferential spinal stenosis, osteoporosis, chronic infections, metal allergies, pregnancy, facet joint arthrosis, inadequate vertebral endplate size, workers' compensation, spinal litigation, body mass index greater than 35, and any isthmic or degenerative spondylolisthesis greater than Grade 1. Eighteen patients fulfilled all follow-up criteria to 24 months. VAS, ODI, presence of back and leg pain, and patient satisfaction scores were recorded. Statistical analysis used simple tests (t-tests for the continuous VAS and ODI scores, and nonparametric sign tests for the back and leg pain ordinal scores). Eight cases had undergone two-level fusion and two cases had undergone 3 level fusion. The remaining cases had undergone single-level fusion. Preoperatively, 75% had persistent back pain and 50% had persistent leg pain, whereas postoperatively 25% had persistent back pain and none had persistent leg pain. ODI and VAS both were improved from pre-op scores at 24 month follow-up (ODI 65.4 +/-1.5 to 29.9 +/- 1.6; VAS 7.7 +/- 0.3 to 3.4 +/- 0.4). Preoperatively, 23% of the patients worked part time and 13% worked full time; these rates increased to 38 and 27%, respectively. Thirty-five percent of the patients remained unemployed. Length and extent of pre-op disability was not described. Preoperatively, 69% never used narcotics and 31% regularly used narcotics and 63% regularly used tramadol; post-op, none used narcotics and 56% occasionally or regularly used tramadol. There were no device or approach related complications. The author concluded, "Analysis of early results indicated that ProDisc lumbar total disc arthroplasty is an efficacious treatment for symptomatic adjacent-segment lumbar discogenic low-back pain following remote fusion."

Also in 2006, Bertagnoli reported on a 104 patient case series for smokers versus nonsmokers, as smoking has been associated with poorer outcomes in spinal fusion surgery (Bertagnoli, Yue et al. 2006b). Patients 18 to 60 years of age were treated with single level disc arthroplasty (L4 to S1) by the primary author. Exclusion criteria included spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, facet arthrosis, inadequate vertebral endplate size, more than one level of spondylosis, neuromuscular disease, pregnancy, workers' compensation, spinal litigation, body mass index greater than 35, and/or any isthmic or degenerative spondylolisthesis greater than Grade 1. Outcomes examined were patient satisfaction, ODI, VAS, and assessments of neurologic, radiographic, and pain medication. A complete radiographic assessment was performed while discography was used only in certain circumstances. Patients were assessed preoperatively and then at 3,6,12, and 24 months. Differential change over time between smokers and nonsmokers was assessed with mixed effects models for continuous variables (ODI and VAS) and generalized estimating equations for patient satisfaction. Power analysis was done as well as a time-smoking interaction check. Only patients with complete data were analyzed. At 2 year follow-up, patient satisfaction was high (87% in nonsmokers, 94% in smokers). Implanted level disc motion ranged from 3 to 7 degrees. ODI average reduction was 10.69 (standard error of the mean of 1.06). Preoperative VAS decreased from 7.5 to 4.5 in smokers and from 7.5 to 3.8 in nonsmokers at the 2 years follow-up. The percentage of patients with leg pain in both groups decreased from about 50% preoperatively to 16% in smokers and 9% in nonsmokers. There was a decrease in medication usage in both smokers and nonsmokers, with preoperative narcotic use being 18% and 16% in smokers and nonsmokers respectively, to 5% and 4% (tramadol – an atypical opioid - use decreased from 27 to 25% in smokers, but increased from 26% to 30% in nonsmokers). The authors stated, "No correlation was determined to exist between clinical outcome and pelvic incidence, tilt, or sacral slope," and, "Complications in this study were primarily limited to those associated with the operative approach and operative field." The author further concluded, "The results of our study indicate that smokers do equally well compared with nonsmokers when ProDisc ADR is used in the treatment of debilitating lumbar spondylosis."

Additionally, Bertagnoli reported in 2006 on 22 patients treated by the primary author with ProDisc, median age of 63 years (range 61-71years) (Bertagnoli, Yue et al. 2006c). Exclusion criteria included spinal stenosis with neurogenic claudication, history of fusion, chronic infections, metal allergies, inadequate vertebral endplate size, pregnancy, workers' compensation recipients, spinal litigation, body mass index greater than 35, and isthmic or degenerative spondylolisthesis greater than Grade 1, significant facet join arthrosis, and patients with T-scores on bone mineral density testing greater than -2.5. There were 17 single-level, four two-level cases and one three-level case of lumbar artificial disc implantation. The primary research questions were whether there was a significant improvement in status from presurgery to 3 months postsurgery and whether there was enduring improvement from 3 months to 2 years postsurgery. Pre-op ODI mean was 27. Post-op ODI mean was 14 at 24 months. VAS pre-op mean score was 8 and the post-op mean score was 4 at 24 month follow-up. Patients in whom bone mineral density was decreased (no details) had concurrent prophylactic vertebroplasty. Of this population, only 15% used narcotics and 40% used tramadol before surgery, with 100% postoperatively not using narcotics after 24 months and 50% not using tramadol (but the reported post-op tramadol usage numbers didn't add up appropriately). Clinical outcomes did not change significantly from the 3 to 24 month interval. Patient satisfaction was high. Overall complication rate in this older age group was higher (2 cases of foot drop, a case of loss of proprioception and vibration sensation necessitating cane assisted ambulation, two cases of subsidence where the T-scores were -2 and -1.76 for a total of 5/20 or 25%). It was stated, "Building on our early experience with two cases of subsidence, we now routinely perform open prophylactic vertebroplasty in which we use 5 to 10 ml of bone cement in the relevant vertebral bodies following implant placement but during the same operative session." The authors did not comment on possible additional changes in biomechanics from vertebroplasty or recent reports in the literature, Trout and Kallmes 2006, of increased adjacent fractures in those who have had vertebroplasty.

In 2006 Siepe reported on 92 patients with a minimum follow-up of 24 months (mean follow-up 34.2 months) (Siepe, Mayer et al. 2006). The objective of the study was to assess functional outcome after total lumbar disc replacement for varying indications, as the authors stated, "Presently, there is no evidence-based consensus on indications or contraindications for TDR." The indications for the procedure included: DDD; DDD with soft disc herniation; osteochondrosis from a previous discectomy; and DDD with Modic changes. The average age of the patients was 42.3 years (range 21.9 – 66.1 years). Operations were performed at 1 (n= 77), 2 (n= 14), and 3 (n= 1) levels. Operating time averaged 115 minutes for one level and 190 minutes for 2 levels. Blood loss averaged 100 ML. VAS preoperative score averaged 7 +/-1.6, with an average 4.2 +/- 2.8 post-op reduction. ODI score averaged 40 +/- 15.6 preoperative and decreased 21 +/- 19 points post-op. Overall, 82.6% of patients were satisfied or highly satisfied. The authors concluded, "that age as a solitary factor does not pose a contraindication to disc replacement", though better functional outcome was observed in younger patients (ages 30 – 40 years). The overall complication rate was 19.6%, requiring revision surgery at the index level in 8.7% of the patients and 2.2% at the non-index level. Complication rate was higher for 2 level replacements (35.7%) versus one level (14.3%). The overall rate for patients returning to their previous job or "some kind of modified professional activity" at their last visit was 68.1%. The authors concluded, "Because of significantly varying outcomes, indications for disc replacement must be defined precisely."

Chung reported on the 2-year clinical and radiographic outcomes of 36 patients by a single surgeon (Chung, Lee et al. 2006). The mean age was 43 years (range 25 to 58 years), with mean follow-up of 37 months. Both one level (25 patients) and two levels (11 patients) were treated. Medication usage preoperatively included nonsteroidal anti-inflammatory medications. Inclusion criteria included minimum disc height of 4mm, ODI of at least 40, and no more than 2 involved levels from L3 to S1. Exclusion criteria included scoliosis, spondylolysis, spondylolithesis, severe facet degeneration and BMD DEXA T score less than -2.5. Also, positive discography was required along with one or more of these findings: vacuum phenomenon, contained herniated nucleus pulposus, high-intensity zone signal, and decrease of intervertebral disc height. Results revealed mean ODI score improvement from 69.2 preoperatively to 34.8 at 6 weeks, 23.0 at 1 year, and 21.0 at 2 years ($p < 0.001$). Mean VAS low back pain scores improved from 7.5 preoperative to 3.7 at 6 weeks, 2.9 at 1 year and 3.0 at 2 years. Mean VAS leg pain scores improved from 4.7 preoperatively to 1.5 at 6 weeks, 1.1 at 1 year, and 1.2 at 2 years ($p < 0.001$). Range of motion at the index level increased from 9.7 degrees preoperatively to 12.7 degrees at 2 years. In an analysis examining factors associated with a successful clinical outcome (defined as > 75 % improvement in ODI), single level versus two level and lower average segmental ROM at 2 years were associated with greater ODI improvement. Two patients had major vein injury which was repaired; three patients had increased leg pain postoperatively that resolved at 6 week follow-up.

Schroven reported on a prospective nonrandomized study of 24 patients (Schroven and Dorofey 2006). Study inclusion included patients between 18 and 60 years of age, 6 months of conservative therapy, and diagnostic CT or MRI. Fourteen patients underwent TDR with ProDisc and 10 patients underwent anterior lumbar intervertebral fusion (ALIF). Follow-up was one year. Statistical analyses were not done due to the small size of the study. Baseline age, gender, spinal level, and ODI (38 out of 60) were comparable. For ProDisc, ODI was 15 at 6 months and 12 at 24 months. In the ALIF group, ODI was 25 at 6 months and 21 at 24 months. Complications in the ProDisc group included one case each of subsidence and facet arthritis. In the ALIF group, one patient had intra-operative hemorrhage. Hospitalization was 3.85 days in the ProDisc group versus 6.3 in the ALIF group. Mean blood loss was 100 ml in the ProDisc group versus 330 ml in the ALIF group. Mean operation time was 1 .5 hours in the ProDisc group versus 2.25 in the ALIF group. The author stated, “The small size of the groups and the limited follow-up period did not allow firm conclusions.”

Preliminary ProDisc randomized trial reports

A 2005 abstract by Delamarter reported on 180 patients in the ProDisc II clinical trial (127 patients underwent TDR and 53 patients had fusion) with follow-up of 2 to 3 years (Delamarter, Zigler et al. 2005). Patient inclusion and exclusion criteria were not reported in this abstract but can be accessed in the FDA summary of safety and effectiveness data. The author noted that with ProDisc TDR can be done at more than one level. He noted improvements in both the VAS and ODI scores were similar for TDR and fusion patients, but that the patient satisfaction was significantly better than the Charite disc (87% v. 73%). No dislocations or device related complications were reported.

Delamarter also reported on “an interim comparative analysis and description of the first 78 randomized patients at 2 years from one site” (Delamarter, Bae et al. 2005). One or two levels of disc disease were included, with evaluation of plain radiographs, MRI, and occasionally discogram/CT scans. Inclusion criteria are listed as: degenerative disc disease in one or two adjacent levels between L3-S1, back and/or leg pain, failure of at least 6 months of conservative therapy, Oswestry score > 20/50 (> 40%), ability to comply with protocol and follow-up, ability to give informed consent, and radiographic evidence of disc degeneration. Exclusion criteria include: more than two levels of degenerative disc disease, endplate dimensions less than 34.5 mm medial-lateral or 27 mm anterior-posterior, known metal and/or polyethylene allergies, prior lumbar fusion surgery, clinically compromised vertebral bodies due to prior trauma, clinically significant degenerative facet disease, lytic spondylolisthesis and/or clinically significant stenosis, degenerative spondylolisthesis > grade 1, back or leg pain of unknown etiology, objective diagnosis of osteoporosis (DEXA scan), presence of metabolic bone disease, morbid obesity (Body Mass Index > 40), pregnancy or expected pregnancy within 3 years, active infection, medications that retard healing (eg. steroids), autoimmune diseases (eg. rheumatoid arthritis), systemic diseases (eg. AIDS, HIV, hepatitis), and active malignancy. Outcomes of ODI and VAS, range of motion, and demographics were analyzed statistically using mixed designs analysis of variance (ANOVA) with repeated measures for assessment interval and a grouping effect for treatment modality (SAS, GLM procedures). Student t-test and Chi squared were used for simple comparisons across treatments. For specific effects, post-hoc pair-wise statistical comparisons were made with Student t-tests or paired t-tests. Graphical comparisons were given for these interim results, with “From 6 months out to 2 years, the disc replacement patients continued to show more improvement than fusion patients, but the difference was not significant. At the longest follow-up, both groups were significantly improved form their preoperative estate.” Delmarter raised an important point in the discussion, that the L5-S1 level is the least mobile in the lumbar spine, and that the difference in sagittal motion in the disc replacement was compared with fusion patients and the difference was not significantly different, and that, “differences in motion are small and harder to detect with a relatively small sample size.”

Zigler reported preliminary results from the clinical study in 3 articles from patients at a single institution (Zigler 2003; Zigler, Burd et al. 2003; Zigler 2004). Zigler 2003 and Zigler, Burd et al. 2003 reported on 39 patients (28 with ProDisc with 6 two-level implants and 11 fusions) and Zigler 2004 reported on 78 patients (55 ProDisc with 25 two-level implants and 23 fusions) with at least a 6 month follow-up. The author noted, “By the end of April 2003, nearly 500 U.S. patients had been implanted as part of the study” (Zigler 2003). Zigler 2004 reported, “When the single-level study arm filled its enrollment in April 2003, continued access was granted for a limited number of single-level cases per month, although all patients were still required to meet the study criteria, and identical data were collected for safety purposes.” Inclusion/exclusion criteria are the same as those listed in Delamarter 2005 except age was also listed as an inclusion criteria (age 18 to 60 years) in Zigler, Burd et al. 2003. Fusion patients were maintained in a corset for 12 weeks. After 3 months, patients were referred to physical therapy, and the brace was discontinued. ProDisc patients were treated in a light corset for 2 weeks and then could be offered some type of physical therapy based on “level of function.” The disc group was active sooner than the fusion group. There was a statistically significant difference in improvement in ODI for the disc group than the fusion group only at 3 months ($p < 0.02$) (bar graph comparison presented). The difference in satisfaction rates were not statistically significant at 6 months ($p = 0.08$). There were no differences in pain scores by VAS (no details). Estimated blood loss (2003 (38 patients): ProDisc, 69 mL versus fusion, 175 mL; 2004 (78 patients): ProDisc, 103 cc, fusion 213 cc) and operative time (2003: ProDisc 75 minutes, versus fusion at 219 minutes; 2004: ProDisc 90 minutes, fusion 232 minutes) were statistically different for the two procedures as was length of hospitalization (2003:2.1 days vs. 3.5 days; 2004: 2.2 days vs. 3.3 days). Adverse events included a reoperation on the second postoperative day for improper device insertion and an iliac vein laceration during the index procedure. One ProDisc patient experienced a deep venous thrombosis and was successfully treated. Four patients in the fusion group complained of graft site pain at the six week follow-up. The authors stated, “No patients in either group had changed their job status following surgery.”

Randomized Controlled Trial

Zigler et al. reported on the prospective, randomized multi-center clinical trial for FDA approval for a single implant level (Zigler, Delamarter et al. 2007). Two hundred ninety-two patients (see below for further identification of number of patients in trial) had surgery between October 2001 and June 2003 at one of 17 investigational sites, implanting 162 investigational subjects and 80 control subjects, and 50 subjects (the first three at each site) enrolled as non-randomized training cases. Control patients received a circumferential fusion consisting of an interbody fusion using a femoral ring allograft and a posterolateral fusion with autogenous iliac crest bone graft, combined with pedicle screw instrumentation. The ProDisc is implanted with an anterior approach.

Inclusion criteria:

- 18 to 60 years old
- Single-level DDD at L3-S1. Diagnosis of DDD requires:
    1. Back and/or leg pain
    2. Radiographic confirmation of any 1 of the following by CT, MRI, discography, plan film, myelography, and/or flexion/extension films:
        1. instability(> 3mm translation of > 5 angulation):
        2. Decreased disc height > 2mm:
        3. Scarring/thickening of annulus fibrosis:
        4. Herniated nucleus pulposus: or
        5. Vacuum phenomenon.
- ODI $\geq$ 40 out of 100
- Failed $\geq$ 6 months of conservative treatment

- Psychosocially. Mentally and physically able to fully comply with protocol, including adhering to follow-up schedule and requirements and filing out forms
- Willing to give written informed consent

Exclusion criteria:

- Number of vertebral levels with DDD > 1
- Patients with involved vertebral endplates dimensionally smaller than 34.5 mm in the medial-lateral and/or 27mm in the anterior-posterior directions
- Known allergy to titanium, polyethylene, cobalt, chromium or molybdenum
- Prior fusion surgery at any vertebral level
- Clinically compromised vertebral bodies at the affected level due to current or past trauma
- Radiographic confirmation of facet joint disease or degeneration
- Lytic spondylolisthesis or spinal stenosis
- Osteoporosis. A screening questionnaire for osteoporosis, SCORE (Simple Calculated Osteoporosis Risk Estimation), will be used to screen patients who require a DEXA bone mineral density measurement. If DEXA is required, exclusion will be defined as a DEXA bone density measure T score ≤ -2.5 (The World Health Organization definition of osteoporosis)
- Back or leg pain of unknown etiology
- Paget's disease, osteomalacia or any other metabolic bone disease (excluding osteoporosis addressed above)
- Degenerative spondylolisthesis of grade > 1
- Morbid obesity defined as a body mass index > 40 or a weight more than 100 lbs. over an ideal body weight
- Pregnant or interested in becoming pregnant over the next 3 years
- Active infection
- Taking any drug known to potentially interfere with bone/soft tissue healing
- Rheumatoid arthritis or other autoimmune disease
- Systemic disease including AIDS, HIV, hepatitis
- Active malignancy

Patient demographics are presented by Zigler et al. for 75 fusion patients and 161 ProDisc patients. P value differences were non-significant for implant level, gender, age, race, body mass index, smoking status and prior surgical treatment. The FDA summary of safety and effectiveness provided data for baseline ODI (fusion N = 80, ProDisc N = 162), baseline VAS pain (fusion N = 78, ProDisc N = 159), and preoperative activity levels that did not show statistically significant differences. The trial design was non-inferiority and the computed power calculation used Blackwelder methodology with the assumption that 85% of patients in both arms would have successful results. The chosen delta in success rates for clinical insignificance was 12.5%. With a type I error of 5% (one sided) and 80% power, the authors stated that the calculation was that of 216 patients, with a 2:1 treatment to control ratio. With a potential dropout rate of 15%, an enrollment of 255 patients was established. A fixed randomization blocking method of 6 assignments per block was used, with random allocations generated in a 2:1 ratio. Clinical and radiographic evaluations were performed at 6 weeks and 3 months (+/- 2 weeks), 6 months (+/- 1 month), and 12, 18, and 24 months (+/- 2 months). These evaluations included ODI, SF-36, VAS pain, physical and neurological exams, and radiographic evaluations, which were used as outcomes for 10 primary endpoints, 6 of which were radiographic. The authors defined the patient to be successful if the individual patient met their criteria for success in all 10 endpoints. The authors noted, "During the course of the study, the FDA required alternative definitions of two of the criteria, ODI and ROM." The authors also stated, "Patient accountability reveals follow-up at 24 months was 98.2%. There was no significant difference at 24 months between the investigational (98.6%) and control (97.1%) groups."

The FDA summary provided, "an account of all subjects enrolled and treated in the study who completed all evaluations at each time point within the windows defined in the approved investigational protocol":

Table 2 - Patient Accountability (FDA Summary)

	Fusion	ProDisc(R)*	ProDisc(NR)**
Enrolled(Preoperative)	93	183	51
Treated	80	162	50
Failures (cumulative at 24mos.)	2	6	0
Expected	78	156	50
Evaluated	71	149	48
Actual	69	142	45
Actual in window	57	124	35
Follow-up rate In window %	73.1%	79.55%	70.0%

* Randomized patients
** Nonrandomized patients

For the definition of overall success with multiple endpoints, the applicant proposed criteria differed from the FDA requested criteria.

Table 3 - Overall Success Components (24 months) (FDA Summary)

	Fusion	ProDisc (R)*	ProDisc(NR)**
ODI ≥ 15% improvement (Applicant proposed criteria)	46/71 64.8%	115/149 77.2%	41/48 85.4%
ODI ≥ 15 point improvement (FDA requested criteria)	39/71 54.9%	101/149 67.8%	36/48 75.0%
Device success	73/75 97.3%	155/161 96.3%	50/50 100%
Neurologic success	57/70 81.4%	135/148 91.2%	40/48 83.3%
Improved SF-36 score	49/70 70.0%	18/149 79.2%	43/48 89.6%
Radiographic success (FDA criteria)	59/69 85.5%	125/143 87.4%	40/45 88.9%
Radiographic success (Applicant criteria)	59/69 85.5%	131/143 91.6%	43/45 95.6%
Overall Success (Applicant proposed criteria)	32/71 45.1%	94/148 63.5%	30/45 66.7%
Overall Success (FDA requested criteria)	29/71 40.8%	79/148 53.4%	25/45 55.6%

The ProDisc mean ODI score dropped more quickly than the fusion group, though at 24 months the difference between the two groups was not statistically significant (FDA Summary) (number of patients measured is not available).

Table 4 - Mean ODI scores (FDA Summary)

	baseline	6wk	3 mos	6mos	12mos	18mos	24mos
ProDisc (randomized)	63.4	41.5	36.4	36.0	35.6	34.7	34.5
Fusion	62.2	49.8	46.6	41.5	40.7	39.8	39.8

For VAS pain scores, where the number of patients measured is not available, the difference between the two groups was not statistically significant at any time point other than 3 months (FDA Summary). VAS satisfaction score difference, where the number of patients measured is not available, was statistically significant at 24 months (ProDisc group 77%, fusion 67%). Mean flexion/extension ROM at 24 months (difference in Cobb measurements) was 0.7 degrees for the fusion group, 7.7 degrees for the ProDisc randomized group, and 8.8 degrees for the ProDisc non-randomized group. Interestingly, there was no statistically significant association between range of motion and success/failure at 24 months.

Adverse events were reported in the FDA Summary. For fusion subjects, 87.5% reported an adverse event, while 84% of the randomized ProDisc subjects and 82% of the nonrandomized subjects reported an adverse event. The number of adverse events considered by the investigator to be device-related was 17% (36/212) in the ProDisc group and 20.0% in the fusion group, with a statistically insignificant difference. Device failures occurred in 6/162 (3.7%) of the randomized ProDisc patients, or, expressed differently, 6/212 (2.8%) of the total ProDisc subjects (4 patients had device anterior migration that necessitated device removal followed by fusion; one subject had revision due to part of the device being inserted backwards; one had fusion for facet disease, with the device left in place). Control subjects experienced device failures in 2/80 (2.5%) of subjects (both had hardware removal subsequent to pain). Clinically significant blood loss (> 1500 cc) was reported in two fusion patients. Three patients developed a deep venous thrombosis (2 ProDisc, 1 fusion). No deaths occurred during the study.

Surgical and hospitalization information provided by the authors included:

- Intra-operative time mean (SD) was 121 (59.2) minutes for ProDisc (N= 160) and 229 (75.9) minutes for fusion (N= 75) patients ($p < 0.0001$).
- Estimated blood loss mean was 204 (231.3) cc for ProDisc (N= 160) and 465 (440) for fusion patients (N= 73) ($p< 0.0001$).
- Length of hospital stay ("determined by the patient's ability to transfer and ambulate independently under oral pain management") mean days was 3.5 (1.29) for ProDisc and 4.4(1.54) for fusion (N= 75) ($p = 0.0001$).

The FDA safety and effectiveness summary reported the above information for 78 to 80 fusion patients and 161-162 ProDisc patients, so the means and standard deviations are different, but the p values are similar. The FDA noted, "While the differences in the means for each of these parameters were statistically significant, in each case, the ranges were similar so the statistical significance may not be clinically significant."

Preoperatively, narcotic usage was 76% in the control group and 84% in the investigational group. In those who did not meet the criteria of overall success, narcotic usage was 76% in the control group and 79% in the investigational group. In those who met the definition of overall success at 24 months, 31% of control and 39% of investigational patients remained on narcotics. More detailed information about medication usage was not provided.

Work status was examined. Pre-operatively, there was no difference in the control (78.1%) and investigational (83.5%) groups in employment. At 24 months, 85.1% of the control group and 92.4% of the investigational group were employed. Detailed information about level of work either pre or post op was not provided.

Synthes also presented this data in the meeting with CMS on January 8, 2007 and discussed how they felt their results were better than the Charite disc results.

4. Medicare Evidence Development and Coverage Advisory Committee (MedCAC) Meeting

The MedCAC was not held for this topic. However, on November 30, 2006 a meeting of the Medicare Coverage Advisory Committee (MCAC) (the predecessor to the MedCAC) was held on the topic of Spinal Fusion for the Treatment of Low Back Pain Secondary to Lumbar Degenerative Disc Disease. The basis for this meeting was established in the CMS decision memorandum for LADR issued on May 16, 2006. In that decision memorandum CMS stated, "There is clearly a tremendous need for additional research on the treatment of degenerative disc disease to include the technology addressed in the NCD—the lumbar artificial disc—and other surgical procedures to include spinal fusion. Therefore, CMS will also convene a Medicare Coverage Advisory Committee at the earliest possible time to address the issue of spinal surgery for degenerative disc disease. We urge the spinal surgery community to discuss the current limitations of the evidence for benefit and to outline the steps needed to develop better evidence." (CMS Decision Memorandum, May 16, 2006)

Since the new LADR technologies were being compared to fusion procedures in the FDA approved IDE trials, CMS determined it was crucial to review the state of the evidence for spinal fusion procedures for the treatment of low back pain due to lumbar DDD.

The TA commissioned by CMS for the MCAC concluded, "The evidence for lumbar spine fusion did not conclusively demonstrate short-term or long-term benefits compared with non-surgical treatment, especially when considering patients over 65 years of age" (McCrory, Turner et al. 2006). Based on the total evidence presented at the meeting, the panel advised CMS that the state of the evidence supporting improved clinical outcomes from lumbar spine fusion for DDD was weak. There was discussion about the need for a clinical trial that would conclusively support the health benefit of lumbar spine fusion for the treatment of DDD.

5. Evidence-based guidelines

No evidence-based guidelines were identified.

6. Public Comments

Public comments sometimes cite the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as the results of individual practitioners or patients are less rigorous and therefore less useful for making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

Initial Public Comment Period

During the initial 30 day public comment period, CMS received 55 comments. Of the 55 comments, 15 commenters supported general coverage for LADR, 26 supported specific coverage of the ProDisc-L total disc replacement, seven supported coverage for the younger, younger-disabled or the under age 60 population (two supported general coverage for the younger-disabled Medicare population, one supported general coverage for patients under age 60, two supported coverage of the ProDisc-L for patients under age 60, one supported coverage of the ProDisc-L for the younger-disabled population, one supported coverage of the ProDisc-L for the younger patients), one supported coverage for ProDisc-L for patients over age 40, one supported future coverage of the Maverick disc, and one commenter supported coverage policy be made at local contractor discretion. Three commenters opposed coverage and one commenter only questioned whether denial to a specific older population was discriminatory. It was noted that five of the comments were essentially duplicate comments and a number of other comments had similar phrases within the substance of the comment. Comments provided by way of a form letter or duplicate text do little to inform our decision.

A. Professional Societies

No comments were received from professional societies during the initial 30 day comment period.

B. General Public Comments

Forty-one of the comments were identified as being from Physicians, of which five were identified has having participated in the FDA approved IDE study for ProDisc®-L total disc replacement. One physician commenter identified himself as the inventor of the ProDisc®-L. Two comments were from patients. Three device manufacturers submitted comments and one comment was from the Blue Cross Blue Shield Association Technology Evaluation Center.

One of the physician commenters who identified himself as the principal investigator for the ProDisc-L trial and first author of the study results stated in reference to the study results, "…many of the study discriminators were patient-derived. To a great degree back pain is subjective, and we have to rely on the patient to tell us of our success. These patients told us that they were more satisfied, had less pain, and experienced improved function with ProDisc-L than with fusion." He further stated, "The treatment of back pain is understandably somewhat controversial. As CMS most recently heard at the MCAC meeting, there are differing opinions on the need for surgical intervention in the treatment of low back pain." He also acknowledged the controversy surrounding disc arthroplasty but did not feel that the controversy supported ignoring or discarding the device. An additional statement by this commenter noted the importance of the procedure being done on the right patient by trained surgeons. The issue of the right patient or carefully selected patients for the success of this procedure was also noted by 20 additional commenters. In addition, this commenter stated, "The design of the disc, with the semi-constrained polyethylene dome and endplates with keels are both design elements that make this disc superior to the Charite." Similar statements reflecting the superiority of the ProDisc-L implant compared to the Charite implant were made by 10 other physician commenters and one of these commenters noted the lower cost of the ProDisc-L implant compared to the Charite.

Some additional general topics in the comments included the importance of rigorous training for surgeons performing the procedure and a number of commenters referred to the impact a CMS decision has in the health community. One physician commenter felt one criterion should be that patients should be over the age of 40 to off set the likelihood of the need for replacement while seven other physician commenters felt coverage appropriate for the under 60 years of age or younger/younger-disabled Medicare population. Additionally, numerous commenters noted that very few Medicare patients would fit the criteria for the procedure, but felt that Medicare should not deny coverage to the few. Some felt this procedure was for the young-active population.

One commenter stated that her husband had the ProDisc surgery in Germany and went from being in constant pain and "90% wheel chair bound" to functioning "as if he never had back problems." A number of comments included references to patients returning to work sooner or leaving the hospital sooner with the LADR procedure compared to a fusion procedure and three mentioned that it might save money.

Of the two patients that commented, one was a patient who had a Charite disc implanted and supported coverage and one was a patient awaiting FDA approval of the Maverick disc so he could have that disc implanted. The patient who was waiting for approval of the Maverick disc stated, "I began researching the availability of ADR in the US and have followed the results of FDA clinical trials closely. …As a result of the studies, I have made the decision to try to hold out for FDA approval and subsequent Medicare authorization of the Maverick disc. The reason for this is because I am 40 years old. The research on polyethylene disc (Charite & Prodisc) is that they will most likely breakdown on me at some point. The Maverick disc, being metal on metal, is expected to last 40 years which would put me at 80."

One commenter felt that the results for LADR were mediocre and that combined with the risk of revision surgery "make the technology questionable". He further stated, "It is difficult to find people enthusiastic about the procedure outside of the industry involved. …Perhaps FDA criteria for approval of new devices ought to be superiority of any new product introduced, rather than equivalence with current technology." One physician commenter who identified himself as a reviewer for an insurance company opposed coverage. He stated, "...THE TWO CASES OF CHARITE LUMBAR DISC REPLACEMENATS [sic] THAT I REVIEWED WERE BOTH DISASTERS. THESE DEVICES SHOULD BE TAKEN OFF THE MARKET UNTIL FURTHER RESEARCH IS DONE."

Three device manufactures submitted comments.

Medtronic (the manufacturer of the Maverick disc) stated in their comment: "…that the primary population for disc arthroplasty is not typically elderly Medicare beneficiaries. However, there may be clinically-appropriate individuals within this population who meet the indications for the procedure and who would receive its resulting clinical benefits. Patient–specific conditions and circumstances, in addition to the clinical evidence supporting use, should be the most important factors in determining which treatments are appropriate and most likely to provide the greatest clinical benefit for an individual beneficiary." They further stated: "[We believe that CMS should allow local Medicare contractors the discretion to develop appropriate local policies, based on the indications for use, evidence supporting the effectiveness of these devices, and other appropriate criteria they develop in collaboration with trained physicians and other healthcare providers in their jurisdictions.]"

DePuy Spine (the manufacturer of the Charite disc) stated in their comment: "We would request the following:

- Give careful consideration to the discussion and conclusions from the recent Medicare Coverage Advisory Committee (MCAC) on Spinal Fusion; particularly being aware that the arthroplasty IDE studies exceeded established minimal standards for health-related disability in the Oswestry Disability Index (ODI) and the Visual Analog Scale (VAS) for pain outcomes.
- Provide coverage to both the CHARITÉ Artificial Disc and ProDisc-L Total Disc Replacement, recognizing that they both meet the standard of "reasonable and necessary".
- Apply the same evidentiary standards for review of the ProDisc-L data as were applied to the CHARITÉ data."

DePuy further commented on the MCAC on Spinal Fusion for the Treatment of Low Back Pain Secondary to Degenerative Disc Disease and stated, "The MCAC addressed concerns that directly pertain to the control arms in all of the lumbar artificial disc trials." They commented on the two validated instruments, Oswestry Disability Index (ODI) and the Visual Analog Scale (VAS), used to measure chronic low back pain that were identified in the TA for the MCAC meeting. They provided a comparison table for the ODI and VAS outcomes for the Charite study and the ProDisc-L study. They further stated, "Since a randomized controlled trial has not been conducted comparing CHARITÉ Artificial Disc and ProDisc-L Total Disc Replacement, conclusion statements directly comparing the two studies with respect to any clinical outcomes are not possible." They went on to identify similarities in the two study designs to support their request. They also noted, "Although both studies excluded patients over the age of 60, the clinical community believes that the clinical benefits for CHARITÉ Artificial Disc and ProDisc-L Total Disc Replacement can be achieved in carefully selected Medicare beneficiaries." They believe this was reinforced by the Spinal Fusion MCAC. DePuy concluded its comment with the following recommendation:

"CMS must determine that the product is reasonable and necessary as a condition of coverage under section 1862(a)(1)(A) of the Social Security Act. The CHARITÉ Artificial Disc and the ProDisc-L Total Disc Replacement meet the criterion for reasonable and necessary for a select patient population as demonstrated in the Level I evidence from the FDA IDE clinical trial results. We further believe that the clinical benefits can be achieved in carefully selected Medicare beneficiaries (including the under 65 disabled population and a more limited number of patients 65 and older) and we strongly support the need for careful patient selection criteria. These patient criteria were detailed in DePuy Spine's previous comment on the NCD.

As CMS considers the merits of this NCD, we request that CMS apply consistent coverage standards when evaluating the levels of evidence for both the ProDisc-L Total Disc Replacement and the CHARITÉ Artificial Disc."

Synthes Spine (the manufacturer of the ProDisc-L disc) stated in their comment: "Surgical treatment of degenerative disc disease is only employed when all other treatment options fail. When all other treatment options are exhausted, the patient must have the choice of whether to pursue surgical intervention and must have access to such treatment should they choose to do so." They provided a summary of the data from their prospective, multi-center, randomized, controlled clinical trial of the ProDisc-L total disc replacement compared to circumferential fusion. The data is more specifically summarized in the evidence section of this document. In reference to that study Synthes stated, "Prodisc-L Total Disc Replacement was proven safe and effective in the treatment of degenerative disc disease as evidenced by FDA's approval of the device on August 14, 2006. More importantly, ProDisc-L Total Disc Replacement was proven to be superior to circumferential fusion in that study." They further stated, "We strongly believe this therapy cannot be labeled experimental or investigational. …The experiment was conducted under the watchful eye of the FDA. The experiment was a success and Prodisc-L proved to be safe and effective in the treatment of degenerative disc disease." and "We further suggest that ProDisc-L meets the criteria of reasonable and necessary. The study, which has been submitted for publication, as well as the clinical experience of physicians both here and abroad as well as the lack of complications, would define the ProDisc-L as reasonable while the need for and the success of the alleviation of chronic pain speaks to the necessity." Finally, Synthes requested the CMS provide coverage for the patients for whom the FDA has indicated it is safe and effective.

D. Comments with Evidence

Four of the commenters provided reference to evidence.

One commenter referenced an article, "The Treatment of Disabling Single level Lumbar Discogenic Low Back Pain in Patients over 60 with Total Disc Arthroplasty Utilizing the Prodisc®Prosthesis: A Prospective Study with 2 Year Minimum Follow-up", Bertagnoli R, Yue JJ, et al., J Neurosurg Spine 4:1-6, 2006. This article was reviewed as part of the evidence review for this proposed decision memorandum.

One commenter submitted an abstract being submitted for a spine meeting. The abstract was about the Maverick™ disc which is another lumbar artificial disc that is not yet approved by the FDA. It was noted that the abstract, Prospective Randomized Series Comparing Maverick™ Lumbar Total Disc Replacement (T.D.R.) with Anterior Lumbar Interbody Fusion (A.L.I.F.) by Pettine and Donner, reported on patients in a prospective randomized controlled trial with an age range of 21 to 55 years. The limited information provided appeared to be results of a RCT from a single site which compared 25 patients receiving the Maverick disc with 11 patients receiving an ALIF utilizing an LT-CAGE with Infuse Bone Graft. The author concluded, "These Maverick T.D.R. results are similar to those reported by six other I.D.E. sites. These combined results of 173 Maverick patients indicate statistical superior clinical outcomes compared to A.L.I.F. at one year and two year follow up." We await publication of the complete data.

The comment provided by the BCBSA Technology Evaluation Center was the attachment of their TEC Report on Artificial Vertebral Disc Replacement from April of 2005 and a link to a article in Forbes: “Dangerous Devices” (Herper and Langreth, Forbes 11/27/06). The TEC report stated “…the use of artificial vertebral discs for degenerative disc disease does not meet the TEC criteria.” (Further summary of the report is provided in Section VII.B.2 of this document.) In the article, “Dangerous Devices”, concerns are expressed about the safety of some of the devices being implanted in people. The article includes the artificial disc as one of the devices of concern. The link to the complete article is http://www.forbes.com/forbes/2006/1127/094_print.html.

Dr. Marnay, the physician who invented the ProDisc-L, submitted a paper titled, NCA Lumbar Artificial Disc Replacement (LADR) ProDisc-L®. This paper has not been published in a peer-reviewed scientific journal. Dr. Marnay disclosed he is under a consulting contract with Synthes Spine (the device manufacturer) and has a financial interest in ProDisc-L®. Dr. Marnay stated: “As inventor of ProDisc-L, using it for almost 17 years with an experience of more than 1000 levels implanted, I strongly believe that it must be offered to the patients under an insurance coverage. The quality of the results demonstrated in all the studies and especially the US FDA IDE, reinforces the conviction that this is a terrific progress in treatment of the **D**egenerative **D**isc **D**isease and this treatment must be delivered to those who need it and are good selected candidates.” He further stated: “The US study design included patients between 18 and 60 years old. However two studies have been conducted (Marnay, Bertagnoli) concerning patients over sixty years old and their results confirm that, following drastic conditions of selection, ProDisc-L should be delivered and reimbursed for this population too.” This paper also references the long history of spinal arthrodesis and the impact of arthrodesis on adjacent levels and facet joints and offers that maintaining mobility through total disc replacement protects adjacent levels from degeneration.

A summary of the paper provides for the following: the importance of the design of the implant; the reproducibility of outcomes from patient to patient; the safety of the implant design and implantation procedure; and, appropriate patient selection as key to the success of any treatment. The paper also provides information about the long-term follow up of a series of patients from 1990-93 showing the safety and maintenance range of motion. The randomized FDA IDE study is referenced and the author stated, “The randomized study so conducted on 286 patients has showed the “Effectiveness and Safety” of this technique in the proper selected patients.” Marnay provided a table of the results of the study of which he states “…shows the quality of the results and the statistically superiority of the ProDisc-L Total disc Replacement on the 360°Fusion.” The FDA ProDisc®-L IDE study results are thoroughly described in the evidence and analysis section of this proposed decision memorandum.

Marnay referenced a comparative study he conducted of Prodisc® Total Disc Replacement in patients below versus over 60 years old single level versus multi levels. This study has not been published in a peer-reviewed scientific journal. This was a prospective study including 476 patients with the same indications as the FDA IDE study but included patients over 60 years of age. Of the 223 patients with single level disc replacement, 12 were over 60 years old (mean age of 64 and range of 60 – 78). Of the 218 patients with multi level disc replacement, 23 were over 60 (mean age 63 and range 60-71). The results of the ODI and VAS were provided in a table breaking out the under age 60 and over age 60 patients by single level and multi level. Marnay stated, “There is no statistically significant difference in the results presented between the two groups, below versus over sixty years old.” Since the data provided from this study is limited and is not presented within the context of a peer-reviewed published article, it is not possible for CMS to fully analyze the results of the study. However, Marnay stated, “Those conclusions about the quality of the results for the patients over 60 treated with a ProDisc-L® procedure are corroborated by the publication of R. Bertagnoli.” and a review of the Bertagnoli publication is included the evidence review and analysis section of this proposed decision memorandum.

In discussing the risk, Marnay poses the following questions, “Do you create the possibility for everyone who needs it and fits with the criteria of selection, to have access to this technology or only let this technology accessible for those who will pay for it? To let the surgeons make what they have always done, use an approved and proven safe technology and test it on the mass of the patients with the quick answer of the “market” or keep on running a slow process of “no accessibility”, accepting not to deliver the best treatment to a mass of population, without objective reasons?” He also discusses how some techniques disappear when better ones come along. Finally Marnay stated, “The reimbursement of the ProDisc-L® procedure is a necessity now. But if we want to make this successful for years and confirm on bigger numbers of patients operated on what we have seen as results, we need to follow the process of selection of the patients already defined (and that is even more fundamental for the patients over 60), to follow the process of trainings of great qualities, to maintain the control of the “good practice of the users” and the permanent studies and publications.”

VIII. Analysis

National coverage determinations (NCDs) are determinations by the Secretary with respect to whether or not a particular item or service is covered nationally under Title XVIII of the Social Security Act §1869(f)(1)(B). In order to be covered by Medicare, an item or service must fall within one or more benefit categories contained within Part A or Part B, and must not be otherwise excluded from coverage. Moreover, with limited exceptions, the expenses incurred for items or services must be “reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member.” § 1862(a)(1)(A). This section presents the agency’s evaluation of the evidence considered and conclusions reached for the assessment questions.

CMS focused on this general question:

Is the evidence sufficient to conclude that LADR with the ProDisc lumbar artificial disc will improve health outcomes in the Medicare population with low back pain due to degenerative disc disease?

Identifying the pain generator in chronic low back pain is challenging due to the anatomic complexity of the spine and poor understanding of neurophysiologic mechanisms of pain sensation (Haldeman 1999). Low back pain symptoms are generally vague. A reliable diagnostic test to determine the exact cause of low back symptoms has yet to be developed, so it is difficult to differentiate between possible sources (MCAC fusion 2006). Treatment of symptoms relies primarily on a subjective measure - clinical judgment. Though the evidence base for conservative treatment is weak, the majority of patients with low back pain will have acceptable results without surgery (MCAC fusion 2006). Persistent low back pain can lead to significant functional limitations. In those who do not respond to non-operative treatment, spinal fusion is an alternative. Spinal fusion for the indication of discogenic pain remains debated with conflicting interpretations of clinical trials even among the experts (MCAC fusion 2006). No universally accepted indication guideline exists to assist the physician in patient management. For patients, improvement in short term pain and function from fusion is unacceptably variable, and long term results remain controversial. In spite of the many ambiguities, the use of spinal fusion surgery in the United States continues to increase (Deyo, Nachemson 2004). The artificial lumbar disc has been developed as an alternative to fusion; however, fusion is used in broad anatomic pathologies, including conditions studied in the Spine Patient Outcomes Research Trial (SPORT): spinal stenosis (a narrowing of the spinal canal), spondylolisthesis (slippage of the spinal unit), and herniated discs. Total disc replacement (TDR) is generally not considered a treatment for degenerative disc disease if these accompanying conditions are present to any significant degree, although a recent article by Siepe et al. stated, “Presently, indications for TDR remain a matter of debate, and an increasing number of reports are available that describe disc replacement procedures for “extended indications” or even commonly accepted contraindications, such as spinal canal stenosis, degenerative scoliosis, residual intersegmental instability following previous fusion, degenerative scoliosis and mobile degenerative spondylolisthesis (Grade I and II)” (Siepe, Mayer et al. 2006).

Current artificial lumbar discs (FDA approved or in development) can vary in material, motion design, or method of fixation. The designs are diverse, and we do not know the effect of the various designs on patient results (Anderson and Rouleau 2004). Additionally, we do not know which surgical protocols influence clinical outcomes (Rousseau, Bradford et al. 2006). As with other discs, published case series predate the U.S. randomized clinical trial reports. These case series studies generally report good results (statistically significant improvements in reported measures), though some reports present limited information. One of the more completely reported case series of single level implantation, Bertagnoli 2005, reported that for 118 subjects at 24 months, 91% of subjects had either occasional pain or no pain, with significant decreases in ODI (53% to 29%) and VAS (7.6 to 3) that were sustained at 24 months (Bertagnoli, Yue et al. 2005). Statistical methods here used a longitudinal approach to assess changes over time. Preoperatively, 83% did not use narcotics, and 90% postoperatively did not use narcotics. Complications included one reoperation. In an older group with the potential for multilevel disease, Bertagnoli 2006 reported on 22 patients with a median age of 63, where vertebroplasty is used for some patients (Bertagnoli, Yue et al. 2006c). In this small case series report that combines two procedures in some patients, no mention is made of possible vertebroplasty-induced changes in biomechanics, or recent reports in the literature, Trout and Kallmes 2006, of increased adjacent fractures in those who have had vertebroplasty. Complications were higher in this older population. Case series data provide weak evidence as the results can be influenced by patient selection, natural history where symptoms wax and wane, or the expectation-response (placebo) effect.

The ProDisc randomized clinical trial has as a comparator 360 degree – circumferential – fusion, a type of fusion that generally has higher successful fusion rates. The trial is designed to demonstrate that the disc is not inferior to this type of fusion; however, it is not clear that a trial designed to demonstrate noninferiority is valid given that the effectiveness of fusion in degenerative disc disease is not well-established in comparison to no treatment (MCAC fusion 2006). Other issues create uncertainty in the trial design. Gotzsche stated that the choice of a clinically relevant difference is, “crucial in noninferiority and equivalence trials for planning the trial, determining sample size, and for interpreting results” (Gotzche 2006); however, there is no justification of the noninferiority margin. The prespecified success rates are significantly different than actual study results – 85% versus 63.5% and 45.1%. It’s not clear why the results were so far off the mark, and again calls in to question if, in fact, noninferiority to fusion also means superiority to conservative management. For a noninferiority comparison, the investigational treatment where the results are not inferior to another treatment is generally considered acceptable if there are other obvious advantages. For the lumbar artificial disc, the advantages are not obvious. Though the disc has been in clinical use in other countries well over 10 years, the design promise of spinal mobility leading to improved outcomes over fusion remains an unproven idea. The available evidence thus far does not provide a direct link between spinal mobility and improved clinical outcomes.

The sponsor defined overall success in the ProDisc trial as a summation of 10 criteria: a 15% improvement in the ODI; lack of adverse events from the device (defined as a device failure requiring revision, re-operation or removal); SF-36 improvement (defined as any numerical improvement); no deterioration in neurological condition; and results of 6 x-rays (quantified as dichotomous outcomes). While there are numerically more measures than in the Charite randomized trial, the requirements are not more stringent. Although pain is the indication for the procedure, measurement of pain is not in these 10 criteria for success. Narcotic use, closely linked to pain, is not reflected in these 10 criteria. Here, as in the Charite trial, a significant number of study participants remained on narcotics at 24 months, even in 39% of ProDisc subjects who met the sponsor's success criteria (details other than certain percentages were not available). The sponsor's proposed ODI criteria are less stringent than the Charite trial and includes patients who lack a clinically significant improvement (15% change of the mean is 9.5 points). It's not clear if simply having any numerical increase in SF-36 has clinical meaning. While a lack of adverse events is a good thing, it is difficult to consider it a health benefit. The radiographic outcomes have uncertain clinical meaning. The authors themselves stated, "Overall success is a mathematical measure and does not directly measure clinical success" (Zigler, Delamarter et al. 2007). They go on to state, that the focus should be on VAS pain, patient satisfaction, ODI, work status, and narcotic usage, "as these parameters may have more clinical relevance." This definition of success is unconvincing as a demonstration of health benefit for an individual patient.

Incomplete patient accountability further complicates interpretation of this study. Zigler 2005, an early report of the IDE trial, reported that 500 patients were enrolled as of March 2003. It is presumed that those not reported in this study are in the two level study, but we don't know. The clinical trial report doesn't list denominators, only percentages, so it's difficult to know who was included in their analysis. The FDA summary does have some denominators. From this summary, we conclude that 9/80 patients are completely excluded from the fusion group, and 12/160 excluded from the ProDisc group (presumed for missing data). For the 24 month results, in excess of these 21 excluded patients, the denominators vary among the outcomes, though the FDA summary stated, "all randomized subjects who completed all evaluations at the 24-month time point, regardless of when the 24-month measurement occurred". In-window protocols are defined, but in the final analysis out of window subjects are included, so it's possible that time could confound these outcomes. When only in-window protocol subjects with FDA criteria are considered in the analysis (which offers a more complete look at the study), the overall success rate was 53.4% for ProDisc and 40.8% for fusion. In the overall success rate, the device failures do not appear to be reported, which is counterintuitive to the definition of failure.

In specific consideration of the Medicare population (who are elderly, disabled, or both), study exclusion criteria of the ProDisc randomized controlled trial limit the generalizability of results. For instance, no one over age 60 was included in the study. The outcomes of those with comorbidities, which could affect the procedure's efficacy and safety, are also unknown. In general, there is limited knowledge of spine operation outcomes in the elderly and those with comorbidities (MCAC fusion 2006). Data were not provided on how many patients were screened to arrive at the enrolled patients. Patients eligible for the lumbar artificial disc implantation, using strict criteria, may be narrowly focused. A study by Huang found that of 100 consecutive patients who had lumbar surgery in one spine surgeon's practice, 95% of patients had one or more contraindications to disc replacement, with the mean number of contraindications of 2.5 per patient (Huang, Sandhu 2004). Perhaps most importantly, "Presently, there is no evidence-based consensus on indications or contraindications for TDR" (Siepe, Mayer 2006).

“Adverse events in spine surgery are often arbitrarily reported as ‘device-related,’ ‘major,’ or ‘preventable,’” and, “These judgments are not always straightforward, and they profoundly influence interpretation of safety data,” (Mirza, Deyo 2006). Adverse events are difficult to judge from the Zigler, Delamarter et al. 2007 paper. In the FDA summary, an adverse event occurred in 84% of ProDisc patients and 87.5% of fusion patients. Device related adverse events occurred in 17.9% of ProDisc patients and 20% of fusion patients. Device problems that required reoperation or revision occurred in 6 ProDisc patients and 2 fusion patients. There are several single case reports of disc failure (Mathew, Blackman et al. 2005; Shim, Lee et al. 2005; Stieber and Donald 2006), but the rate of these occurrences can not presently be judged.

Bertagnoli noted, “Most of the complications in total disc replacement procedures are iatrogenic; wrong indications, poor implantation technique, and improper positioning of the implant are the most likely causes. Isolated device-related complications are rare (e.g., subsidence, body fractures, polyethylene extrusion, and problems due to polyethylene wear). Due to stringently controlled inclusion groups, small study populations, and lack of long-term follow-up, only limited data are available. Lessons learned from hip and knee arthroplasty, however, suggest that the incidence of complications increases with duration of follow-up” (Bertagnoli, Zigler et al. 2005).

In the particular circumstance of Medicare patients, generally, there is limited knowledge of spine operation complications in the elderly and those with comorbidities (MCAC fusion 2006). Information on safety in real-world patient situations is essential for informed choice.

The published abstract by Marnay on over 7 years of follow-up of ProDisc I implants gives some evidence of long term viability, as does report by Tropiano 2005, which seems to be the same study group (Marnay 2002; Tropiano Huang et al. 2005). Marnay reported a reduction in VAS and mean post operative ODI score. Tropiano reported improvement in other measures but does not report VAS or ODI. It was unclear if follow-up was systematic or when the measures listed were recorded. Though complications were reported, it was not clear how many patients were free from complications at follow-up. There was no correlation of mobility with outcome.

The major premise of spine segmental motion preservation is that adjacent level disease is increased by fusion surgery and that motion preservation will prevent this. However, the available literature does not provide evidence of a direct link between spinal mobility and improved clinical outcomes to support this premise. As Hassett reported in his study on aging and the nonoperated lumbar spine, there was radiographic evidence of progressing osteoarthritis of 3% to 4% per year, without symptom correlation (Hassett, Hart et al. 2003). This rate has also been quoted as the risk of adjacent level disease after fusion. Hilibrand and Robins stated, “However, based on the present scientific literature, it is still unclear whether these radiographic and clinical findings are the result of the spinal fusion with the iatrogenic production of a rigid motion segment of whether these represent the progression of the natural history of the underlying degenerative disease” (Hilibrand and Robins 2004).

In a publication of Roundtables in Spine Surgery, Roh et al. noted, “Although the advent of modern total disc designs and implantation techniques have heralded a recent paradigm shift in the treatment of symptomatic lumbar disc degeneration, strict adherence to proper patient selection and perfect implant placement must be upheld by all spine surgeons involved in total disc replacement. In light of the recent popularity of this surgical procedure, however careful consideration of surgical strategies to revise these implants will undoubtedly be required in the not-so distant future” (Roh, Pappou et al. 2005). Considering long term use of the device and the possibility of revision, patients need to be assured that should the device become unacceptable, the treatment will not be worse than this non-life threatening disease.

Conclusion

Chronic back pain from degenerative disc disease is complex and can be difficult to treat. The LADR trials use the fusion procedure as a control in a noninferiority trial. As was presented in the CMS MCAC on spine fusion for DDD, “The evidence for lumbar spine fusion did not conclusively demonstrate short-term or long-term benefits compared with non-surgical treatment, especially when considering patients over 65 years of age” (McCrory, Turner et al. 2006). As far as the studies on LADR, many questions remain regarding selection criteria, adverse events, and long term outcomes for spine surgery in general. The ProDisc lumbar artificial disc FDA IDE clinical trial limited patients to the ages between 18 to 60 years old, as with the previous randomized controlled disc trial on the Charite disc, continuing to exclude the age group with the highest prevalence of degenerative disc disease. In addition, the limited information regarding the Maverick™ lumbar artificial disc trial provided through public comment identified the patient population as being up to 55 years of age. Based on the age limitations set in the FDA IDE trials for lumbar artificial discs thus far, CMS is convinced that the indications for LADR will exclude the over age 60 population and that these age limitations are not specific to one manufacture’s disc implant. Therefore, the reconsideration of the NCD will address the procedure of LADR rather than LADR with a specific manufacture’s disc implant. The one small case series on patients over 60 years of age was insufficient to draw any reasonable and necessary conclusion of the benefit provided to the population.

Due to the lack of evidence of benefit for those Medicare beneficiaries over the age of 60, CMS proposes to noncover LADR in this population. Some evidence does exist for patients age 60 and under. However, rather than confirm the results of earlier case series studies, the ProDisc FDA IDE noninferiority clinical trial creates more uncertainty in benefit due to certain issues including trial design and reporting. Patient studies without a comparison group make it difficult to draw clear conclusion on the benefit of treatment, though some individual patients with this poorly understood, potentially disabling problem may benefit. In consideration of the difficulty in arriving at a clear conclusion of the benefit of this technology for the Medicare beneficiary 60 years of age and under population, no change is proposed and current coverage will continue to be determined by local contractors.

IX. Decision

The Centers for Medicare and Medicaid Services (CMS) proposes that LADR is not reasonable and necessary for the Medicare population over sixty years of age. Therefore, CMS is proposing that Section 150.1 of the Medicare NCD Manual be amended to reflect the proposed change from non-coverage for a specific LADR implant to non-coverage for the LADR procedure for the Medicare population over sixty years of age. For Medicare beneficiaries sixty years of age and under, there is no national coverage determination, leaving such determinations to be made on a local basis.

Appendices [PDF, 110KB]
Back to Top

Bibliography

Anderson P, Rouleau J. Intervertebral disc arthroplasty. Spine, 2004;29:2779-2786.

Bertagnoli R, Kumar S. Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. Eur Spine J 2002;11:S131-S136.

Bertagnoli R, Yue, J, Fenk-Mayer A, et al. The treatment of disabling single-level lumbar discogenic low back pain with total disc arthroplasty utilizing the ProDisc prosthesis: a prospective study with 2 -year minimum follow-up. Spine. 2005a Oct 1; 30(19):2230-6.

Bertagnoli R, Yue J, Shah R, et al. The Treatment of Disabling Multilevel Lumbar Discogenic Low Back Pain With Total Disc Arthroplasty Utilizing the ProDisc Prosthesis A Prospective Study With 2- Year Minimum Follow-up. Spine 2005b: 30(19);2192-2199.

Bertagnoli R, Zigler J, Karg A, et al. Complications and strategies for revision surgery in total disc replacement. Orthop Clin N Am 2005;36:389-395.

Bertagnoli R, Yue J, Fenk-Mayer, et al. Treatment of symptomatic adjacent-segment degeneration after lumbar fusion with total disc arthroplasty by using the ProDisc prosthesis: a prospective study with 2- year minimum follow up. J Neurosurg Spine 2006a Feb;4:91-97.

Bertagnoli R, Yue J, Kershaw T et al. Lumbar total disc arthroplasty utilizing the ProDisc prosthesis in smokers versus nonsmokers: a prospective study with 2-year minimum follow-up. Spine 2006b; 31 (9):992-7.

Bertagnoli R, Yue J, Nanieva R, et al. Lumbar total disc arthroplasty in patients older than 60 years of age: a prospective study of the ProDisc prosthesis with 2-year minimum follow-up period. J Neurosurg Spine, 2006c Feb; 4(2):85-90.

Blue Cross and Blue Shield technology assessment. Artificial vertebral disc replacement April 2005. http://www.bcbs.com/betterknowledge/tec/vols/20/20_01.pdf. Accessed 3/07/07.

BCBS TEC MAP March 2007 web site. http://www.bcbs.com/betterknowledge/tec/press/actions-taken-by-the-medical-1.html. Accessed 3/08/07.

Boden D, Davis D, Dina T, et al. Abnormal magnetic resonance scans of the lumbar spine in asymptomatic subjects. J Bone Joint Surg 1990;72-A:403-8.

Chung SS, Lee CS, Kang CS. Lumbar total disc replacement using ProDisc II: A prospective study with a 2-year minimum follow-up. J Spinal Disord Tech 2006; 19(6):411-5.

CMS Decision Memorandum. Lumbar Artificial Disc Replacement. May 16, 2006. http://www.cms.hhs.gov/mcd/viewdecisionmemo.asp?id=170. Accessed 3/08/07.

CMS MCAC. Spinal Fusion for the Treatment of Low Back Pain Secondary to Lumbar Degenerative Disc Disease. Transcript -November 30,2006. http://www.cms.hhs.gov/faca/downloads/id37e.pdf. Accessed 3/08/07

Delamarter RB, Zigler J, Ohnmeiss D, et al. Lumbar total disc replacement: a 2 to 3 year report from 2 centers in the United States clinical trial for the Prodisc-II prosthesis. The Spine Journal 2005;5(1S):189S.

Delamarter RB, Bae HW, Pradhan BB. Clinical results of ProDisc-II lumbar total disc replacement: report from the United States clinical trial. Orthop Clin North Am 2005; 36 (3):301-13.

Deyo RA, Nachemson A, Mirza SK. Spinal –Fusion Surgery – The Case for Restraint. N Engl J M ed Feb 2004;350 (7):722-726.

FDA guidance document for the preparation of IDEs for spinal systems. January 13, 2000. http://www.fda.gov/cdrh/ode/87.pdf. Accessed 1-6-06.

FDA Summary of Safety and Effectiveness Data for Expedited Premarket Approval (PMA) August 14, 2006. http://www.fda.gov/cdrh/pdf5/p050010b.pdf. Accessed 3-01-07.

Gibson JNA, Waddell G. Surgery for degenerative lumbar spondylosis. The Cochrane Database of Systematic Reviews 2005, Issue 2. Art. No.: CD001352.pub2. DOI:10.1002/14651858.CD001352.pub2.

Gotzche P. Lessons From and Cautions About Noninferiority and Equivalence Randomized Trials. JAMA March 2006;295(10):1172-1174.

Gunzburg R, Mayer HM, Szpalski M, et al. Arthroplasty of the spine: the long quest for mobility. Eur Spine J 2002;11:S63 -S64.

Hägg O, Fritzell P, Nordwall A: The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J, 2003 12:12–20.

Haldeman S. Low back pain. Current physiologic concepts. Neurol Clin 1999;17:1-15.

Hassett G, Hart DJ, Manek N, et al. Risk factors for progression of lumbar spine disc degeneration: The Chingford Study. Arthrit Rheum 2003;48:3112-7.

Henanff AL, Giraudeau B, Baron G, et al. Quality of Reporting of Noninferiority and Equivalence Randomized Trials. JAMA March 2006;295(10):1147-1151.

Hilibrand AS, Robbins M. Adjacent segment degeneration and adjacent segment disease: the consequences of spinal fusion? Spine J 2004; 4 (6 Suppl):190S-4S.

Huang R, Sandhu H. The current status of lumbar total disc replacement. Orthop Clin N Am 2004;35: 33-42.

Krismer M. Fusion of the lumbar spine. A consideration of the indications. J Bone Joint Surg (Br.) 2002;84:783-794.

Manek NJ, MacGregor AJ. Epidemiology of back disorders: prevalence, risk factors, and prognosis. Curr Opin Rheumatol 2005;17:134-140.

Marnay T. Lumbar disc replacement, 7-11 years results with Prodisc. Eur Spine J 2002; 11(suppl 1):S19.

Mathew P, Blackman M, Redla S et al. Bilateral pedicle fractures following anterior dislocation of the polyethylene inlay of a ProDisc artificial disc replacement: a case report of an unusual complication. Spine 2005; 30(11):E311-4.

Mayer HM, Wiechert K, Korge A, et al. Minimally invasive total disc replacement: surgical technique and preliminary clinical results. Eur Spine J 2002; 11 Suppl 2:S124-30.

McCrory D, Patwardhan M, Richardson W, et al. AHRQ Draft Technology Assessment, Spinal Fusion for Treatment of Degenerative Disease Affecting the Lumbar Spine Draft. October 23, 2006. http://www.cms.hhs.gov/determinationprocess/downloads/id41ta.pdf. Accessed 3/07/07.

Mizra SK, Deyo RA, Heagerty PJ, et al. Towards standardized measurement of adverse events in spine surgery: conceptual model and pilot evaluation. BMC Musculoskeletal Disorders 2006.7:53.

Piaggio G, Elbourne D, Altman D, et al. Reporting of Noninferiority and Equivalence Randomized Trials; An Extension of the CONSORT Statement. JAMA March 2006;295(10):1152-1160.

Roh JS, Pappou IP, Girardi FP, Cammisa FP. The ProDisc-L : FDA Studies and European Experience. McAfee PC, Geisler FH, Scott-Young M, editors. Roundtables In Spine Surgery. Missouri: Quality Medical Publishing, Inc.; Volume 1, 2; 2005. p.117-132.

Rousseau MA, Bradford DS, Bertagnoli R, et al. Disc arthroplasty design influences intervertebral kinematics and facet forces. The Spine Journal 2006; 6:258-266.

Salovey P, Seiber WJ, Smith AF, et al. Reporting chronic pain episodes on health surveys. National Center for Health Statistics. Vital Health Stat 6(6). 1992. http://www.cdc.gov/nchs/data/series/sr_06/sr06_006.pdf. Accessed 3/08/2007.

Santos E, Polly D, Mehbod A, et al. Disc arthroplasty: lessons learned from total joint arthroplasty. Spin JH 2004:4:182S-189S.

Schroven I, Dorofey D. Vertical split fracture of the vertebral body following total disc replacement using ProDisc: report of two cases. Acta Orthop Belg. 2006:72:83-6.

Shim CS, Lee S, Maeng DH et al. Vertical split fracture of the vertebral body following total disc replacement using ProDisc: report of two cases. J Spinal Disord Tech 2005; 18(5):465-9.

Siepe CJ, Mayer HM, Wiechert K et al. Clinical results of total lumbar disc replacement with ProDisc II: three-year results for different indications. Spine 2006; 31(17):1923-32.

Sott AH, Harrison DJ. Increasing age does not affect good outcome after lumbar disc replacement. International Orthopedics (SICOT) 2000;24:50-53.

Stieber JR, Donald GD, 3rd. Early failure of lumbar disc replacement: case report and review of the literature. J Spinal Disord Tech 2006; 19(1):55-60.

Tropiano P, Huang RC, Girardi FP et al. Lumbar disc replacement: preliminary results with ProDisc II after a minimum follow-up period of 1 year. J Spinal Disord Tech 2003; 16(4):362-8.

Tropiano P, Huang RC, Girardi FP et al. Lumbar total disc replacement. Seven to eleven-year follow-up. J Bone Joint Surg Am 2005; 87(3):490-6.

Trout AT, Kallmes DF, Kaufmann TJ. New Fractures after Vertevro0plasty: Adjacent Fractures Occur Significantly Sooner. Am J Neuroradiol 2006;27:217-23.

U.S. Preventive Services Task Force. Guide to clinical preventive services, 2nd ed. Baltimore: Williams & Wilkins, 1996: 699.

Zigler JE. Clinical results with ProDisc: European experience and U.S. investigation device exemption study. Spine 2003; 28 (20):S163-6.

Zigler JE, Burd TA, Vialle EN et al. Lumbar spine arthroplasty: early results using the ProDisc II: a prospective randomized trial of arthroplasty versus fusion. J Spinal Disord Tech 2003; 16(4):352-61..

Zigler JE. Lumbar spine arthroplasty using the ProDisc II. Spine J 2004; 4 (6 Suppl):260S-7S.

Zigler J, Delamarter RB, Sapivak JM, et al. Results of the Prospective, Randomized, Multi-Center FDA Investigational Device Exemption Study of the ProDisc®-L Total Disc Replacement versus Circumferential Fusion for the Treatment of One Level Degenerative Disc Disease. Spine 2007; 31(11):1155-1162.

To submit comments, please use the orange "comment" button at the top of the page.

Back to Top