Artificial Lumbar Disc Replacement

Assessment
Program
Volume 22, No. 2
June 2007

Executive Summary

Background
Low back pain is an extremely common condition that affects up to 20% of the U.S. population. Degenerative disc disease (DDD) is a major common cause of low back pain. Replacing the disc with an artificial disc is one proposed method to alleviate the pain associated with DDD. There are two U.S Food and Drug Administration- (FDA-) approved products, the Charité disc and the ProDisc.

Objective
This Assessment will review the available evidence to determine if artificial lumbar disc replacement is an effective treatment for chronic degenerative disc disease.

Search Strategy
MEDLINE search (via PubMed) through May 2007, using the search terms "Charité" or "ProDisc" or "intervertebral disk/[MeSH] OR spinal OR spine OR lumbar."

Selection Criteria
Central focus of the report is the randomized clinical trials of the Charité and ProDisc artificial discs. Case series reports were abstracted, but are relatively uninformative regarding effectiveness because of lack of control groups.

Main Results
Each disc type has been evaluated with one randomized, clinical trial. The trials were designed as noninferiority trials, with the comparator being fusion. The validity of a noninferiority trial rests on several assumptions:

- The comparator treatment should have well-known and precise knowledge of effectiveness compared to no treatment. This knowledge and the noninferiority margin designated for the trial should assure that the new treatment is superior to no treatment.
- The trial should also achieve historical levels of effectiveness in the known comparator.
- Finally, an acceptable margin of inferiority is reasonable for a new treatment if there are obvious advantages of the new treatment, such as patient acceptability, convenience, invasiveness, or cost.

The effectiveness of fusion for chronic degenerative disc disease is not well established. There are few clinical trials and results are inconsistent. Neither of the studies discussed the effectiveness of fusion or justified the size of the noninferiority margin. The possible advantages of the artificial disc in terms of physical functioning should be measurable as a principal outcome.

BlueCross BlueShield Association

An Association of Independent Blue Cross and Blue Shield Plans

KAISER PERMANENTE®

The clinical trial of the Charité disc randomized 205 patients to the Charité disc and 99 patients to receive fusion using the BAK cage. Patients receiving the Charité disc had a composite success rate of 57.1%, and patients receiving the BAK cage had a success rate of 46.5%. This met the specified noninferiority criteria with a p-value of 0.0001. This would not have met statistical significance for a traditional superiority trial with conventional levels of significance. The overall success rate was lower than specified in pretrial sample size calculations. In other outcomes, the Charité disc was not statistically significantly superior to BAK fusion, with the exception of patient satisfaction.

The clinical trial of the ProDisc randomized 162 patients to the ProDisc and 80 patients to circumferential fusion. The overall success rate was 63.5% for the ProDisc, and 45.1% for fusion. This met statistical significance using traditional p-value calculation for a superiority trial. The success rates were lower than specified in pretrial sample size calculations. However, success rates were based on patients completing the trial rather than the full, randomized sample. The lack of documentation of missing data and subjects makes it difficult to know which patients were ultimately included in the analysis, particularly those with early device failures, of which there were 6 ProDisc device failures and 2 fusion device failures. Some outcomes showed the ProDisc to be superior, others showed no difference from fusion.

Author's Conclusions and Comments

Given what is known about fusion as a comparator treatment, both noninferiority trials may not provide evidence of efficacy. The specific noninferiority margins are not justified. The lower-than-expected success rates also raise additional questions regarding the validity of a noninferiority trial and the noninferiority margin selected. Viewed from the perspective of superiority trials, both trials are also suspect. The Charité trial showed little evidence of superiority, and the ProDisc analysis is problematic because of missing values and uncertain outcomes for all patients.

Based on the available evidence, the Blue Cross and Blue Shield Medical Advisory Panel made the following judgments about whether the artificial lumbar disc for treatment of degenerative disc disease meets the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) criteria.

1. **The technology must have final approval from the appropriate governmental regulatory bodies.**

In October 2004, the U.S. Food and Drug Administration (FDA) granted premarket application (PMA) approval for the Charité Artificial Disc, stating that the device is indicated for spinal arthroplasty in skeletally mature patients with DDD at one level from L4-S1. DDD is defined as discogenic back pain with degeneration of the disc confirmed by patient history and radiographic studies. As a condition of approval, the manufacturer has agreed to conduct a postapproval study, using a maximum of 366 patients (201 randomized investigational subjects, 67 training investigational subjects, and 98 control subjects). Postapproval study patients will be evaluated for a period of 5 years post-implantation.

The FDA granted approval to the ProDisc in August 2006 and is indicated for spinal arthroplasty in skeletally mature patients with DDD at one level from L3-S1. Patients should have no more than Grade 1 spondylolisthesis at the involved level. Patients receiving the ProDisc should have failed at least 6 months of conservative treatment prior to implantation. The manufacturer agreed to conduct a postapproval study to obtain 5-year follow-up data from all subjects in the clinical study. The study will utilize the same endpoints as the Investigational Device Exemption clinical study, and also evaluate adjacent segment degeneration and correlation of range-of-motion with Oswestry Disability Index and visual analog scale scores.

2. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.

Current evidence supporting the effectiveness of artificial lumbar disc is insufficient. Case series evidence is inadequate to establish efficacy. The randomized trials of the Charité disc and ProDisc are suspect as valid noninferiority trials and do not prove superiority.

3. The technology must improve the net health outcome; and
4. The technology must be as beneficial as any established alternatives.

The evidence is insufficient to determine whether the use of artificial lumbar discs improves the net health outcome or whether they are as beneficial as any established alternatives.

5. The improvement must be attainable outside the investigational settings.

Whether the use of artificial lumbar discs improves health outcomes has not been established in the investigational settings.

Therefore, the use of artificial lumbar discs for degenerative disc disease does not meet the TEC criteria.

Contents

Published in cooperation with Kaiser Foundation Health Plan and Southern California Permanente Medical Group.

TEC Staff Contributors

Author—David H. Mark, M.D., M.P.H.; **TEC Executive Director**—Naomi Aronson, Ph.D.; **Managing Scientific Editor**—Kathleen M. Ziegler, Pharm.D.; **Research/Editorial Staff**—Claudia J. Bonnell, B.S.N., M.L.S.; Maxine A. Gere, M.S.

Assessment Objective

Low back pain is an extremely common condition that affects up to 20% of the U.S. population. Degenerative disc disease (DDD) is a major common cause of low back pain. Replacing the disc with an artificial disc is one proposed method to alleviate the pain associated with DDD. Note that the artificial disc has specific selection criteria that restrict its use to a subset of patients with DDD. Its comparator, lumbar fusion, has much broader indications. This Assessment will review the available evidence to determine if disc replacement is an effective treatment for chronic degenerative disc disease.

Background

Low Back Pain

Low back pain is an extremely common condition, with an incidence approaching 15% to 20% in the U.S. (Deyo and Tsui-Wu 1987). The physiologic basis for low back pain is highly complex, in large part because the lumbar spine itself is an unusually complex anatomic structure. The spine is the only organ comprised of bones, joints, ligaments, fatty tissue, multiple layers of muscles, and nerves (including peripheral nerves, nerve roots, sensory ganglia, autonomic ganglia, and the spinal cord). Furthermore, these structures are supplied by an intricate arterial and venous system and lie in close proximity to the skin with its sensory receptors. Diagnosis and treatment of low back pain require an understanding of many different types of tissue, as well as knowledge of the biomechanics of complex spinal structures, the manner in which they can be injured, and the variety of biochemical manners in which each of these structures responds to trauma and to aging. In addition, certain other psychosocial factors that affect the manner in which pain is processed centrally in the brain should be considered (Haldeman 1999).

Spinal structures and tissues that possess either unmyelinated nerve innervation or documentable substance P or related peptides are assumed to have the capacity to cause pain. Such structures include the posterior facet joints, bones and periosteum, muscles, tendons, fascia, ligaments, nerve roots, dorsal root ganglia, dura mater and the intervertebral disc (Haldeman 1999).

The Vertebral Disc as a Source of Low Back Pain

Disc pain is a potential cause of low back pain. There is a lack of consensus in the medical literature as to what extent the intervertebral disc is innervated (Pope and DeVocht 1999). Once believed to be inert because nerve endings could not be demonstrated in the nucleus or inner annular fibers, the intervertebral disc is now known to contain fine nerve endings in the outer one-third of the annulus. These nerve endings are immunoreactive to a number of pain-related neuropeptides (substance P, calcitonin-gene-related peptide [CGRP], and vasoactive intestinal peptide [VIP]) (Weinstein et al. 1988). Impulses from these free nerve endings in the outer third of the disc and the adjacent longitudinal ligaments reach the spinal cord through a number of sensory nerves in the following manner:

- the posterior and posterolateral annulus, together with the posterior longitudinal ligament and the ventral dura, is innervated by the sinu-vertebral nerves;
- the anterior and lateral aspect of the disc, together with the periosteum of the vertebral bodies, is innervated through the gray rami communicantes.

The sinu-vertebral nerves have been shown to innervate tissues one or two layers above or below their origin, a finding that may explain the poor localization of lumbar pain.

Evidence that these nerve endings observed in the outer one-third of the disc may be a source of low back pain is based upon several clinical observations. First, it has been demonstrated in human volunteers that injection of 11% sodium chloride into the intervertebral disc causes, after a few seconds, very severe pain with deep aching across the back and poor localization (Hirsch et al. 1963). There is an early case report describing a patient who had low back pain produced by pulling on a nylon suture that was looped through the intervertebral disc (Smyth and Wright 1958).

More recently, Kuslich and Ahern (1994) observed that 33–40% of patients in their large back surgery series had significant pain when the affected central or lateral annulus was stimulated. Finally, other investigators have reported that examination of pathologic discs reveals unusually profuse innervation (Yoshizawa et al. 1980). Grigg and colleagues

(1986) suggest from evidence in animal studies that the intervertebral disc contains a relatively rich supply of what they termed “silent nociceptors”—nerve endings that are not readily excited by mechanical stress, but which, when exposed to pain-inducing substances accompanying inflammatory, degenerative, or traumatic processes, become exquisitely responsive (Grigg et al. 1986).

Types of Intervertebral Disc Damage

Different types of disc injuries can potentially lead to pain. These include annular tears, disc protrusions with extrusion of nucleus pulposus into radial tears in the annular fiber of the disc, and disc herniation, in which nucleus pulposus tissue escapes the confines of the annulus. These events cause pain by stretching or tearing peripheral innervated disc fibers or by generating an irritating inflammatory reaction in adjacent spinal tissues (Swenson 1999).

Degenerative changes in the collagen fibers of the intervertebral disc may also lead to increased focal segment instability. As the intervertebral disc ages, nuclear hydrostatic pressure is lost, leading to buckling of the annular lamellae, increased shear stress across the annular wall, and eventually annular delamination and fissuring of the annular wall. All of these changes have been shown to alter disc mechanics, making annular disruption, a precursor of disc herniation, more likely. Degenerative disc disease is considered to result from the inability of the disc’s reparative capacity to keep pace with the trauma that occurs with activities of daily living.

Degenerative disc disease can be accompanied by spinal stenosis, a narrowing of the spinal column which causes nerve compression, and spondylolisthesis, a displaced vertebral disc. These conditions cause additional symptoms in addition to back pain. Total disc replacement is not considered a treatment for DDD accompanied by these other conditions; thus this report will not discuss further these conditions.

Alternative Treatments for Low Back Pain Due to Degenerative Disc Disease

Gibson et al. (1999) in a Cochrane review of surgery for low back pain, distinguishes between treatments directed specifically at disc prolapse (herniation) versus treatments directed at degenerative lumbar spondylosis (same as DDD). Total disc replacement is intended only to treat DDD but without accompanying spinal stenosis or severe spondylolisthesis. Thus, treatments for disc prolapse such as chymopapain and various techniques for discectomy are not alternative treatments to total disc replacement. Treatments for spinal stenosis or spondylolisthesis such as decompression with or without fusion are also not alternative treatments to total disc replacement.

Nonsurgical therapies are the first-line treatment for back pain associated with DDD. Such treatments include physical therapy, massage, and manipulation. Some controlled trials have shown these modalities to be effective (Cherkin et al. 2003). However, many patients do not respond to such treatments.

Intradiscal electrothermal treatment (IDET) is another alternative treatment (TEC Assessments, Vol. 18, No. 19; 2004). Stronger evidence of its efficacy is still needed, as there are few rigorous randomized, clinical trials evaluating this treatment.

Surgical arthrodesis, or fusion, is considered the current standard surgical treatment for DDD which is not responsive to other treatments. Elimination of motion across the disc space and reduction of loads on disc tissues theoretically result in pain relief. Evidence supporting the efficacy of fusion is relatively sparse.

A review of spinal fusion surgery by Deyo et al. (2004) found that national survey data shows that use of spinal fusion has increased rapidly (i.e., annual numbers of procedures increasing by 77% from 1996 to 2001), owing to new technological advances (e.g., bone morphogenetic protein), financial incentives, and controversial expansion of indications (e.g., discogenic back pain without evidence of sciatica), and a high rate of reoperations. The authors state that, “Fundamental problems plague the study of spinal fusion, including the lack of definitive methods to confirm a solid fusion, a weak association between solid fusion and pain relief, and the placebo effect of surgery for pain relief.” They further state that, “Evidence-based practice for degenerative spine disorders might reserve the use of spinal fusions for spondylolisthesis and only rare cases of disk herniation or spinal stenosis without spondylolisthesis,” and that “More evidence from clinical trials should be required for degenerative disk disease to be an accepted indication” (Deyo et al. 2004). Regarding the use of “emerging spinal implants,” such as artificial discs, the review states that, “If ongoing trials suggest results equivalent to those of spinal fusion, it may be faint praise, given the paucity of evidence that spinal fusion is safe and effective for common indications” (Deyo et al. 2004).

A 1992 review by Turner et al. could find no randomized trials of fusion. Combining many studies of fusion performed for many different clinical indications, they found an average of 68% of patients reported a satisfactory outcome. A 1999 Cochrane review (Gibson et al.) concluded that at that time there was no acceptable evidence of any form of fusion for degenerative lumbar spondylosis, back pain, or “instability.” The authors could find no randomized clinical trials comparing fusion to a nonsurgical alternative, only trials which compared surgical techniques of fusion to each other.

Since the Cochrane review, there have been 2 published clinical trials comparing lumbar fusion to a nonsurgical alternative treatment for patients with chronic back pain due to DDD. Fritzell et al. (2001) conducted a multicenter randomized controlled trial comparing 3 techniques of lumbar fusion to nonsurgical treatment. Enrollment criteria included patients with chronic pain, severe disability, pain attributed to DDD, and no neurologic compromise due to herniated disc, spondylolisthesis, or spinal stenosis. There was no specified nonsurgical treatment, but it was described as commonly used physical therapies. Overall results of the trial are shown in Table 1. Patients receiving fusion therapy reported mean changes from baseline in the range of 18-33% for various pain and disability scores, whereas patients in the control group had changes ranging from -21 to 8% for the same outcomes, all between-group changes being statistically significant. In terms of patients’ overall assessment, 63% of patients receiving fusion reported being better or much better, compared to 29% of control patients. Critics of the study have pointed to the modest effect of surgery (up to 30% mean score change), and the fact that control patients may not have received optimal nonsurgical treatment (Deyo et al. 2004).

The other randomized trial by Ivar Brox et al. (2003) assigned a specific cognitive and exercise regimen to the nonsurgical patients. Enrollment criteria for this study were roughly similar to the other clinical trial, and outcomes were assessed at 1 year. In this study, patients receiving fusion reported improvements ranging from 36 to 49% on pain and disability scales, but patients in the control arm also reported similar improvements in these scores, resulting in differences which were not statistically significant for most outcomes. Although this trial was much smaller than the study by Fritzell et al. (2001), the point estimates of effect for each arm are very similar to each other, and confidence intervals sufficiently narrow to rule out a large clinical benefit of surgery. The authors believe that the difference in results between the two studies is caused by the specific intervention used in the nonsurgical group, which produced improvements similar in magnitude to the surgical fusion group.

The relative sparseness of controlled clinical trial data regarding the effectiveness of fusion for DDD makes uncertain the validity of it as a valid comparator to total disc replacement. It cannot be ruled out that some of the improvements associated with lumbar fusion are due to natural history, placebo effects, or co-interventions such as rehabilitation and exercise programs. It would be difficult to compare retrospective studies of fusion with case series results of artificial disc because of the very restrictive selection criteria for the artificial disc. Complicating the evaluation of fusion is the variety of techniques and devices used to perform the procedure. Pedicle screws and intervertebral fusion cages are two types of devices implanted during some procedures. Clinical trial results comparing use of these devices have not produced consistent results.

Common complications of fusion include instrument failure (7%), complications at the bone donor site (11%), neural injuries (3%), and failure to achieve a solid fusion or pseudarthrosis (15%) (Deyo et al. 2004). Fusion is thought to cause increased rate of disc degeneration in spinal segments adjacent to the fusion.

The clinical trial of the Charité disc compared the disc versus BAK fusion. There is no clinical trial evidence comparing the BAK fusion device to other methods of fusion. The clinical trial which led to FDA approval was a single arm trial enrolling 947 patients, with comparison to literature-based controls who received fusion by a similar (anterior or posterior) approach. In the single-arm analysis, total success as assessed by 1) successful fusion, 2) improvement in pain, 3) maintenance or improvement in function, and 4) maintenance

Table 1. Trials of Lumbar Fusion Versus Nonsurgical Treatment

Study	Outcomes of Fusion			Outcomes of Control			Between-group p-value
Fritzell et al. (2001)	2 years			2 years			
Surgery n=201		pre	post		pre	post	
Control n=63	Back pain	64.2	43.2		62.6	58.3	0.0002
	Leg pain	35.3	29.0		35.6	42.6	0.005
	Oswestry	47.3	35.7		48.4	45.6	0.015
	MVAS	63.7	45.6		65.5	60.4	0.004
	GFS	49.1	34.1		47.6	45.5	0.005
Ivar Brox et al. (2003)	1 year			1 year			
Surgery n=35		pre	post		pre	post	
Control n=26	Back pain	62.1	39.4		64.1	48.7	0.14
	Leg pain	43.5	26.6		34.0	35.5	0.002
	Oswestry	42.0	26.4		43.0	29.7	0.33
	GFS	35.9	18.3		44.6	22.6	0.50

GFS General Function Score
MVAS Million Visual Analog Score

or improvement in muscle strength, was achieved by 72% of patients (184/254) at 24 months. Despite differences in patient selection and outcome assessment, outcomes of BAK fusion were reported to be roughly similar to literature controls, but this comparison lacked matched controls or other methods to control for differences in patient selection and outcome assessment.

The clinical trial of the ProDisc compared the disc to circumferential fusion. Circumferential fusion is a more extensive procedure than either anterior or posterior fusion, but, in general, produces higher successful fusion rates. It is sometimes used selectively in patients with risk factors for unsuccessful fusion. Some randomized studies have shown better clinical outcomes with circumferential fusion than with postolateral fusion, but it is uncertain if these results would generalize to patients eligible for the artificial disc (Christensen et al. 2002; Videbaek et al. 2006). At least one study suggested equivalent outcomes between anterior fusion and circumferential fusion (Madan and Boeree 2003).

Outcome Assessment

Outcomes of treatments for back pain have been compared using a variety of techniques. Most common are pain scales measured on a visual analog scale. Various questionnaires have been developed to additionally capture measures of physical functioning. One of the more common measurement scales in use specific to patients with back pain is the Oswestry Disability Index (ODI), originally developed in 1976. The validity, consistency, and reproducibility of the ODI were extensively reviewed by Roland and Fairbank (2000). This review cites an article by Meade et al. (1986), which suggests that a 4-point difference in the ODI is the minimum difference carrying clinical significance. This article (Roland and Fairbank 2000) also cites a personal communication from the FDA, which states that the FDA has chosen a minimum 15-point change in spinal surgery patients as a clinically meaningful difference in the ODI.

Total Disc Replacement

Prosthetic discs are meant to replace the function of the native disc. Such discs are meant to perform the function of a natural disc, including range of motion and transmission and absorption of compressive loads. There are several disc replacement products that have been developed or are in development and testing. They include Charité, ProDisc, Maverick, and Flexicore. All discs consist of metal endplates that affix to the vertebral bones and some mechanism in between that allows for motion in several planes. The Charité artificial disc uses a polyethylene core which can shift dynamically within the disc space during spinal motion. The ProDisc also uses a polyethylene core. This Assessment will only evaluate the evidence on the Charité disc and the ProDisc as they are the only products to have FDA approval at this time. None of the other disc replacements have more rigorous evidence of efficacy.

The Charité disc was developed in the 1980s and has been used more extensively than other artificial discs. It has been estimated that the disc has been implanted in more than 5,000 patients since 1987. During this early period of use, no randomized, clinical trials had been published. Only results from case series had been published. Through this clinical experience the clinical indications and contraindications for use of the disc were developed. It is indicated only on a subset of patients for whom fusion is often performed. The disc is indicated in individuals with degenerative disc disease at a single level at either L4-5 or L5-S1. It is contraindicated in individuals with instability (caused by spondylolisthesis, fracture, or tumor), osteoporosis, prior major spine surgery, facet joint arthritis, and spinal infection.

The ProDisc was first implanted by Dr. Theirry Marnay in 1990. It has been used in Europe for many years. Throughout its period of use in Europe, there have been no randomized, clinical trials. Several case series have been published and will be reviewed in the evidence section of this report.

Beyond the evaluation of the disc as an effective treatment for back pain due to DDD, the disc raises unique concerns. According to de Kleuver et al. (2002), who did a systematic review of disc replacement, these issues are 1) loosening, 2) subsidence (migration of prosthesis), 3) polyethylene wear, 4) mobility, 5) adjacent disc degeneration, 6) complication rate, and 7) salvage procedures in case of failure. In their review, de Kleuver et al. found no mention of loosening, subsidence, and polyethylene wear. Not all papers addressed the mobility provided by the disc. The operated segment appears to move with a reported average range of motion of 5–12 degrees, but mobility of the segment is frequently lost due to the need for subsequent surgical fusion or spontaneous fusion. None of the studies address whether disc replacement can reduce the rate of adjacent segment degeneration, which is considered to be a problem of fusion. The complication rate was highly variable. For 411 patients in the studies considered in the review, there were no infections, 8 vascular injuries, and 6 thrombotic complications. Regarding a safe salvage procedure, several studies report doing a posterolateral fusion.

There are several published case reports of particular types of complications of either the Charité disc or the ProDisc. Van Ooij et al. (2007) describe 4 patients in which severe wear of the polyethylene cores of the Charité disc caused failure of the disc. Mathew et al. (2005), Shim et al. (2005), and Stieber and Donald (2006) all published single case reports of particular types of acute failure of the ProDisc. However, these case reports do not allow inference regarding the incidence rate of failure, and whether such failure rates are higher or lower than would be achieved with alternative treatments.

The manufacturer of the Charité disc sponsored a randomized, clinical trial comparing the disc to anterior fusion using the BAK fusion cage. The study has been published in the peer reviewed literature (Blumenthal et al. 2005). Prepublication results and a detailed statistical critique by an U.S. Food and Drug Administration (FDA) statistician were also published by the FDA. Geisler et al. (2004) reported selected findings from this clinical trial, but the numbers differ slightly from the FDA document, and the focus of the publication was the neurological complications. The results of this clinical trial led the FDA to approve the Charité disc in October 2004 (see "FDA Status"). A detailed description and assessment of this trial will be reported in the Review of Evidence section of this report.

As a condition of approval, the manufacturer agreed to conduct a postapproval study, using a maximum of 366 patients (201 randomized investigational subjects, 67 training investigational subjects, and 98 control subjects) (U.S. Food and Drug Administration 2004c). Postapproval study patients will be evaluated for a period of 5 years post-implantation. The primary endpoint of the study will evaluate "overall success," defined as:

- improvement of at least 15 points in the ODI score compared to baseline;
- no device failures requiring revision, reoperation, or removal;
- absence of major complications, defined as major vessel injury, or major neurological deterioration (e.g., nerve root injury); and
- maintenance or improvement in neurological status versus baseline, with no permanent neurological deficits compared to baseline status.

The manufacturer of the ProDisc also sponsored a randomized, clinical trial. This study has just been published (Zigler et al. 2007), and information and findings are also available from the "summary of safety and effectiveness data" available from the FDA. The ProDisc was approved by the FDA on the basis of information provided by the company, and no advisory committee was convened to review the data.

FDA Status. In October 2004, the FDA granted premarket application (PMA) approval for the Charité Artificial Disc, stating that the device is indicated for spinal arthroplasty in skeletally mature patients with DDD at one level from L4-S1. DDD is defined as discogenic back pain with degeneration of the disc confirmed by patient history and radiographic studies. As a condition of approval, the manufacturer has agreed to conduct a postapproval study, using a maximum of 366 patients (201 randomized investigational subjects, 67 training investigational subjects, and 98 control subjects). Postapproval study patients will be evaluated for a period of 5 years post-implantation.

The FDA granted approval to the ProDisc in August 2006 and is indicated for spinal arthroplasty in skeletally mature patients with DDD at one level from L3-S1. Patients should have no more than Grade 1 spondylolisthesis at the involved level. Patients receiving the ProDisc should have failed at least 6 months of conservative treatment prior to implantation. The manufacturer agreed to conduct a postapproval study to obtain 5-year follow-up data from all subjects in the clinical study. The study will utilize the same endpoints as the IDE clinical study, and also evaluate adjacent segment degeneration and correlation of range-of-motion with ODI and VAS scores.

Methods

Search Methods

The MEDLINE database (via PubMed) was searched from 1980 through May 2007, using the search terms "Charité" or "ProDisc" or "intervertebral disk/[MeSH®] OR spinal OR spine OR lumbar." The search was limited to English-language citations involving human subjects. In addition, bibliographies of key articles were reviewed for relevant citations.

Study Selection

The randomized, controlled trial of the Charité disc is published in the peer-reviewed literature, with slightly different data than found in FDA materials due to more complete follow-up. From other literature found, studies of consecutive case series of artificial disc procedures using the Charité disc were included if they reported patient outcomes. The randomized, controlled trial of the ProDisc has been published in the peer-reviewed literature, with essentially identical numbers to the FDA material. However, the FDA material has more thorough reporting of certain information.

Medical Advisory Panel Review

This Assessment was reviewed by the Blue Cross and Blue Shield Association Medical Advisory Panel (MAP) on February 21, 2007. In order to maintain the timeliness of the scientific information in this Assessment, literature searches were performed subsequent to the Panel's review (see "Search Methods"). If the search updates identified any additional studies that met the criteria for detailed review, the results of these studies were included in the tables and text where appropriate. There were no studies that would change the conclusion of this Assessment.

Formulation of the Assessment

Patient Indications

Potential candidates for artificial disc replacement have chronic low back pain attributed to degenerative disc disease, lack of improvement with nonoperative treatment, and none of the contraindications for the procedure, which include multilevel disease, spinal stenosis or spondylolisthesis, scoliosis, previous major spine surgery, neurologic symptoms and other minor contraindications. These contraindications make artificial disc replacement an option for a subset of patients for whom fusion is indicated. Patients who require procedures in addition to fusion, such as laminectomy and/or decompression are not candidates for the Charité disc or the ProDisc.

Technologies to be Compared

The current surgical alternative is fusion. Comparison to fusion is complicated by the uncertain efficacy of fusion, as documented in the background section of this report. Nonsurgical alternatives include a variety of cognitive, behavioral, exercise, manipulation, and physical therapy approaches.

Health Outcomes

Benefits. The benefits of treatment for low back pain include pain relief and restoration of function (increased mobility and flexibility, enhanced exercise or sitting tolerance, and return to work). The mobility and flexibility of the artificial disc could potentially improve physical functioning and also may reduce adjacent spinal segment degeneration, but these outcomes have not been specifically assessed.

Harms. Potential harms that could occur after artificial disc replacement include worsened symptoms and complications due to the procedure. However, these harms may also occur with surgical fusion.

Assessment Questions

Does artificial lumbar disc replacement using the Charité disc improve health outcomes in terms of pain relief and restoration of function among patients with chronic discogenic low back pain, compared with fusion or other treatments?

Does artificial lumbar disc replacement using the ProDisc disc improve health outcomes in terms of pain relief and restoration of function among patients with chronic discogenic low back pain, compared with fusion or other treatments?

Review of Evidence

Charité Disc

Case Series Reports

Results from case series reports of disc replacement surgery are shown in Table 2. It is not really possible to make inferences regarding the success of the procedure in these types of studies because there is no control

Table 2. Patient Outcomes of Charité Artificial Disc as Reported in Case Series

Study	Sample Size	Back Pain Outcomes	Complications/Poor Outcomes
Cinotti et al. (1996)	46	2 year outcomes self rating 24% excellent 39% good 30% fair 7% poor	8 pts eventually had fusion 1 bilateral radiculopathy 1 anterior dislocation/reop 4 spontaneous fusion 0 loosened prostheses
David (1993)	22	1 year outcomes self rating 65% exc/good	0 migration 0 infections 1 case of L5 sciatica requiring removal 1 case dislocated prosthesis
Lemaire et al. (1997)	105	Mean follow-up 51 months % with good improvement 79%	5 vascular complications 2 temporary neurologic sx 4 cases bone complications 0 loosening 0 migration of prosthesis
Griffith et al. (1994)	93	Mean follow-up 1 year Average pain reduction 6 preop to 3 postop	1 device failure 5 migration of prosthesis 1 dislocation 30 other procedural complication 3 patients with reoperations
Sott and Harrison (2000)	14	Mean follow-up 48 months Good outcome 10/14 Fair outcome 2/14 Poor outcome 2/14 (good- >75% pain relief, return to work, ≤ slight physical restriction, no analgesics)	1 migration-asymptomatic
Zeegers et al. (1999)	46	2-year outcomes 65% improved low back pain 64% improved leg pain 81% return to work 83% "no regret"	2 not properly positioned no significant migration 7 reoperations for complications

group. Differences in patient selection criteria may confound assessment of the procedure. However, these case reports might provide generalizable information regarding complication rates and may provide useful information on the longitudinal trajectory of outcomes among patients receiving the procedure. None of the 6 case series reported in the table report rigorous preoperative and postoperative pain or functional scales, reporting postoperative states only. A variety of complications are reported, including vascular and neurologic outcomes. Migration of the prosthesis is reported in several of the studies. It is difficult to calculate rates of any particular complication because of the varied format of reporting between studies.

The review by de Kleuver et al. (2003) included a few additional case series reported in foreign-language journals. They did not find any comparative trials. In addition to the lack of any controlled comparative studies, they noted high rates of secondary arthrodesis either due to spontaneous bony bridging or to the need for subsequent surgical arthrodesis. They concluded that there was insufficient data to assess the performance of disc replacement.

Charité Clinical Trial

The prior TEC Assessment reviewing this clinical trial used FDA documents to report the results of this trial. When the FDA reviewed and critiqued the trial, some patients had not reached 24 months of follow-up or had missed their 24-month visits. It appears that the major difference between the FDA documents and the peer-reviewed publication (Blumenthal et al. 2005) is the inclusion of these additional patients to the results. The results as reported in the peer-reviewed publication will be shown.

The stated purpose of the clinical trial was to investigate the safety and effectiveness of the Charité artificial disc compared to the BAK Interbody Fusion Device for the treatment of single-level degenerative disc disease. Patients were to be randomized in a 2:1 ratio and treated in 15 different centers. Outcomes were to be assessed at regular intervals out to 24 months. At each site, 5 patients were to receive the Charité disc before patients were randomized.

Methods and Statistical Analysis Plan. The principal outcome to be assessed in the study was a composite outcome, where success was achieved if all the following were found:

1. Improvement in the Oswestry Disability Index =25% at 24 months compared to baseline.
2. No device failures requiring revision, reoperation or removal.
3. Absence of major complications, defined as major blood vessel injury, neurologic damage, or nerve root injury.
4. Maintenance or improvement in neurological status at 24 months, with no new permanent neurological deficits compared to baseline.

Other outcomes included work status, visual analog scale rating of pain, SF-36 scores, and adverse events.

The study was designed as a noninferiority trial with a d=0.15, which means that the 95% confidence interval of the difference between the Charité disc and BAK fusion cage should exclude a success rate of the Charité disc that is 15% worse than BAK fusion. For the trial sample size, the Charité disc could have a success rate that was up to 4.9% worse than BAK fusion and still meet noninferiority criterion. In the trial publication, p-values for the composite outcome were based on a noninferiority calculation that rejects the hypothesis that Charité disc is inferior. However, p-values calculated for all other outcomes, including the separate components of the composite outcome, appear to use the traditional p-value calculation.

A trial that is designed and analyzed as a noninferiority trial usually establishes a less-stringent standard for demonstrating efficacy than a standard clinical trial. Such trials are often employed when there is some margin of acceptable inferiority of a new technology in its principal outcome that is offset by some other advantage, such as less morbidity, less invasiveness, better acceptability to patients, or lower cost. In the case of the Charité disc, there are no major offsetting advantages that are immediately evident or proven, as it is simply proposed to provide greater relief of back pain. The Charité disc might provide greater flexibility than conventional fusion, but there is no firm evidence to show this. Theoretically, the rotation allowed by the artificial disc might prevent subsequent complications in adjacent discs.

In addition, a noninferiority trial assumes well-established and precise knowledge of the effectiveness of the comparator treatment, in this case, fusion. Knowledge of the effectiveness of the comparator is critical to establishing a valid noninferiority margin, so that if the trial should meet noninferiority criteria, it is certain that the new treatment is superior to placebo treatment. As was reviewed earlier in this Assessment, such information is lacking for fusion.

The inclusion and exclusion criteria were fairly typical of the type of patients for which this artificial disc or fusion alone is indicated. Patients had symptomatic degenerative disc disease at a single level unresponsive to conservative treatment, but did not have nerve root compression, spinal stenosis, or severe spondylolisthesis. Patients could have had prior discectomy, laminotomy, laminectomy, or nucleolysis at the same level.

Clinical Trial Results. Approximately 300 patients were randomized and analyzed (Table 3). In prepublication FDA documents, subjects who discontinued the study early were included in the analysis but considered to be failures. The peer-reviewed publication does not explicitly state this, but in comparing the two documents, it appears that patients who discontinued the study are considered failures. The peer review publication also differs from the FDA documents in that additional patients ultimately had 24-month follow-up evaluations.

Patients receiving the artificial disc had an overall composite success rate of 57.1%, and patients receiving the BAK cage had a success rate of 46.5%. This met the specified noninferiority (p=0.0001) (Table 4). However, the success rates were much lower than the hypothesized 70% success rate. A lower-than-expected success rate raises concerns regarding the validity of a noninferiority trial, and whether the new treatment under consideration would show superiority to placebo (Kaul and Diamond 2006).

Although the Charité disc had a higher success rate than the BAK cage, this difference would not have met traditional significance levels for a superiority trial. Analyses of the separate components of the composite outcome showed that the difference was mostly attributable to the improvement in the Oswestry disability score, with a success rate of 64% for the artificial disc and 50.5% for the BAK cage (p=0.0338). However, in comparing mean improvement between groups among completers, there was no statistically significant difference (Charité 48.5% improvement, control 42.4% improvement, p=0.267). Since patients discontinuing the trial were imputed as failures and a higher proportion of fusion patients discontinued, the success rates based on imputation may be biased in favor of the Charité disc. The pain visual analog scale and SF-36 scores were not significantly different between groups. Patient satisfaction showed a higher satisfaction rates with the Charité disc (73% vs. 55%, p=0.009). However, patients' subjective satisfaction could be affected by the fact that the study was unblinded.

In addition to adverse events that were incorporated into the composite outcome, other adverse events occurring in all 205 randomized patients were tabulated (Table 5). The category "severe or life-threatening events" is apparently a much broader category of events than the "major complications" used in the composite outcome, and "device-related adverse events" does not correspond to any part of the composite outcome. As shown in the table, the artificial disc had higher rates of severe and life-threatening events (15% versus 9%) and device-related adverse events (7.3% versus 4%), but lower rates of device failure (4.9% versus 8.1%). Statistical significance of these differences is not reported, but none of these differences is statistically significant.

FDA In-Depth Statistical Review. The FDA performed an in-depth statistical review of the prepublication data in order to expedite the premarket application process. The review made several important comments regarding the conduct and analysis of the trial. The review noted that there was no statistical analysis plan in the original protocol documents. A statistical analysis plan appears to have been finalized at a date by which most trial data were probably available; thus the sponsor needs to clarify when the statistical analysis plan was finalized and whether the analysis plan was developed or modified based on preliminary review of the data.

The FDA reviewer wanted more detail on the sensitivity analyses performed to account for missing observations, particularly more specific data on the "last value carried forward" technique of data imputation (such sensitivity analyses were not reported in the peer-

Table 3. Clinical Trial, Population Characteristics, and Patient Follow-up Characteristics

	Charité Disc	BAK Cage
Enrolled and randomized	205	99
Completers	176	74
Failures and deaths	13	8
Early discontinuation*	16	17
Characteristics of Population		
Men %	46%	55%
Mean age	39.5	40.1
Level L4L5 operated	29%	33%
Level L5S1 operated	71%	67%

* In prepublication FDA documents, early discontinuations were imputed to be failures, in peer-reviewed documents, nothing is stated.

Table 4. Results of Clinical Trial at 24-month Endpoint

	Charité Disc		BAK Cage		p-value
Composite outcome (ITT)	57.1%	(117/205)	46.5%	(46/99)	noninferior*
Composite outcome (completers only)	63.6%	(115/177)	56.8%	(46/81)	noninferior*
Components of Composite Outcome					
>25% improvement Oswestry scale	63.9%	(131/205)	50.5%	(50/99)	0.0338
Lack of device failure	94.6%	(194/205)	92%	(91/99)	0.448**
Lack of major complication	99%	(203/205)	99%	(98/99)	1.00**
Lack of neurologic deterioration	88%	(160/182)	87%	(74/85)	0.44**
Other Outcomes					
Mean % reduction Oswestry (completers only)	48.5%		42.4%		0.267
Pain Visual Analog Scale Significant improvement	74%	(128/182)	62%	(49/85)	0.076
SF-36					
Physical component, >15% improved	73%		66%		0.348
Mental component, >15% improved	50%		55%		0.496
Patient satisfaction, % satisfied	73%		55%		0.009

* Not statistically significant by traditional "superiority" p-value, significant p-value to rule out greater than 15% inferior success rate with Charité disc.
** As abstracted from paper, cannot reproduce these calculations

Table 5. Results of Clinical Trial, Adverse Events (Randomized Population)

	Charité Disc		BAK Cage	
Patients with severe or life-threatening adverse events	15%	(30/205)	9%	(9/99)
Device-related adverse events	7.3%	(15/205)	4%	(4/99)
Device failures	4.9%	(10/205)	8.1%	(8/99)

reviewed publication). The reviewer calculated some scenarios imputing missing data, and found that an analysis of a "worst-case" scenario favoring BAK fusion, imputing success for BAK fusion missing observations and imputing failure for artificial disc missing observations, did not meet noninferiority criteria. Imputation of discontinued patients as failures favored the artificial disc group, as there were fewer such patients in the artificial disc group. Finally, the reviewer found a few calculation errors, and suggested alternative modeling techniques to account for imbalances between groups.

Summary of Charité Disc

In sum, the evidence supporting the effectiveness of the Charité artificial disc is limited. Case series provide little evidence of efficacy, because without randomized controls, outcomes can be influenced by patient selection, placebo effects, or natural history.

The clinical trial of the Charité artificial disc has several potential issues affecting a straight forward interpretation of its results. The analysis showed noninferiority compared to BAK fusion, but did not show superiority. A noninferiority criterion usually implies some trade-off in the principal outcomes for some other tangible trade-off. However, there is no immediately evident advantage to use of the artificial disc. The point estimate of 57% success does not show the artificial disc to be a highly successful treatment.

ProDisc

Case Series Reports

Results from 6 case series reports and one small nonrandomized comparison study are shown in the table (Table 6). It is not really possible to make inferences regarding the success of the procedure in these types of studies because there is no control group. Differences in patient selection criteria may confound assessment of the procedure. However, these case reports might provide generalizable information regarding complication rates and may provide useful information on the longitudinal trajectory of outcomes among patients receiving the procedure. As compared to the case series reports of the Charité disc, reporting of back pain has become more uniform, with 5 of 6 case series reporting preoperative and follow-up Oswestry Disability Scores.

ProDisc Clinical Trial

The clinical trial evaluating the ProDisc is in many ways similar to the Charité clinical trial (Tables 7–9). Patients were randomized in a 2:1 ratio to receive either the ProDisc or a circumferential fusion. A circumferential fusion consists of an anterior interbody fusion with femoral ring allograft and a postero-lateral fusion with iliac bone graft in combination with pedicle screws.

Methods and Statistical Analysis Plan. The principal outcome to be assessed in the study was a composite outcome where success was determined if all the following were achieved:

1. Improvement in the Oswestry Disability Index =15% (or by FDA criterion of 15 points).
2. No device failures requiring revision, reoperation or removal.
3. Any improvement in SF-36 beyond baseline
4. Maintenance or improvement in neurological status.
5. Several postoperative radiographic findings consistent with operative success.

Compared to the success criteria of the Charité clinical trial, the minimum improvement in ODI required for success is less stringent, and radiographic and SF-36 criteria have been added.

The study was also designed as a noninferiority trial with a d=0.125, which means that the confidence interval of the difference in success rates must rule out that the ProDisc is 12.5% worse than fusion. (The FDA requested a 10% noninferiority margin.) The sample size calculation assumed that 85% of patients in both groups would have a successful result. The same concerns about the validity of a noninferiority trial that were mentioned in the description of the Charité trial apply to this trial.

The inclusion and exclusion criteria were fairly similar to those of the Charité clinical trial. Of the potentially notable differences, the ProDisc trial required a more severe minimum level of back disability as measured by the ODI (>40 versus >30). Patients with nerve root compression or a positive straight leg raising test were excluded from the Charité clinical trial. The ProDisc trial included patients with L3-L4 disc disease, whereas the Charité trial excluded this level.

Table 6. Case Series Studies of ProDisc

Study	n Analyzed	Follow-up %	Follow-up Interval	Preop Status		Follow-up Status		Pre-post Significance
Tropiano et al. (2005)	55	86	mean 8.7 yrs	Low-back pain (0-3 points)	2.73	Low-back pain (0-3 points)	1.35	<0.0001
				Lower limb pain (0-3 points)	2.42	Lower limb pain (0-3 points)	0.67	<0.0001
				Impairment (0-3 points)	2.02	Impairment (0-3 points)	0.78	<0.0001
				Stauffer-Coventry Score (0-20)	7.04	Stauffer-Coventry Score (0-20)	16.1	<0.0001
Bertagnoli et al. (2006)	70 nonsmokers	95% overall	2 years	VAS pain	7.5	VAS pain	3.8	<0.001
				ODI	52	ODI	32	<0.001
				% leg pain	48.6	% leg pain	9	<0.001
	34 smokers			VAS pain	7.5	VAS pain	4.5	<0.001
				ODI	55	ODI	28	<0.001
				% leg pain	50	% leg pain	16	<0.001
Bertagnoli et al. (2006a)	25 multi-level disc disease	86	2 years	VAS pain	8.3	VAS pain	2.1	<0.001
				ODI	65	ODI	21.6	<0.001
				% regular pain	92	% regular pain	8	<0.001
Bertagnoli et al. (2006b)	22 all >60 yrs old	100	2 years	VAS pain	8.0	VAS pain	4.0	<0.001
				ODI (raw score)	27.3	ODI (raw score)	14.5	<0.001
Chung et al. (2006)	36	95	mean 3 yrs	VAS back pain	7.5	VAS back pain	3.0	Sig
				VAS leg pain	4.7	VAS leg pain	1.2	Sig
				ODI	69.2	ODI	21	<0.001
Siepe et al. (2006)	92	98	34 mth mean 24 mth min	VAS pain	7	VAS pain	3.8	<0.001
				ODI	39.9	ODI	18.9	<0.001
Nonrandomized Comparison Study								
Schroven et al. (2006)	14 ProDisc	100	1 year	ODI	38.4	ODI	12.5	NR
	10 fusion			ODI	38	ODI	21.4	

Table 7. ProDisc Trial Patient Characteristics

	Fusion	ProDisc	p-value
Gender [n (%)]			
Male	34 (45.3%)	82 (50.9%)	0.48
Female	41 (54.7%)	79 (49.1%)	
Mean age in years (SD)	40.4 (7.6)	38.7 (8.0)	0.13
Race [n (%)]			
Caucasian	59 (78.7%)	133 (82.6%)	0.629
African-American	5 (6.7%)	5 (3.1%)	
Hispanic	10 (13.3%)	18 (11.2%)	
Asian-American	0 (0.0%)	2 (1.2%)	
Other	1 (1.3%)	3 (1.9%)	
Mean body mass index (kg/m^2) (SD)	27.3 (4.3)	26.7 (4.2)	0.48
Smoking Status			
Never	34 (45.3%)	87 (54.0%)	0.1995
Former	17 (22.7%)	40 (24.8%)	
Current	24 (32.0%)	34 (21.1%)	
Any prior surgical treatment	23 (30.7%)	57 (35.4%)	0.56
Baseline ODI (SD)	62.7 (10.3)	63.4 (12.6)	0.61

Table 8. Outcomes of ProDisc Trial at 24 Months

	Fusion	n Analyzed	ProDisc-L	n Analyzed	p-value
Components of Overall Success					
ODI 15% improvement	64.8%	71	77.2%	149	0.039
ODI 15-point improvement	54.9%	71	67.8%	149	0.0449
Device Success	97.3%	75	96.3%	161	1
Neurological success	81.4%	70	91.2%	148	0.0341
SF-36	70.0%	70	79.2%	149	0.0943
Radiographic success	85.5%	69	91.6%	143	NR
Radiographic success (FDA criteria)	85.5%	69	87.4%	143	NR
Overall Success	45.1%	71	63.5%	148	0.0053
Overall success (FDA criteria)	40.8%	71	53.4%	148	0.044
Other Outcomes					
Mean ODI scores	39.8	NR	34.5	NR	NS
VAS pain scale	32	NR	37	NR	0.08
VAS patient satisfaction	67.3	NR	76.7	NR	0.015

Table 9. Adverse Events of ProDisc

	Fusion		ProDisc	
Type of Event	% (n) patients	Events	% (n) patients	Events
All adverse events	87.5% (70)	256	84% (136)	136
All device-related adverse events	20% (16)	34	17.9% (29)	50
Device failures	2.7% (2)	2	3.7% (6)	6

Clinical Trial Results. According to the FDA Summary, 162 patients received the ProDisc and 80 patients received fusion. This differs from the published report (Zigler et al. 2007) that reports 161 ProDisc patients and 75 fusion patients. The prepublication draft does not report any denominators in results tables, but appears to report identical results (in terms of percentages or rates) to the FDA Summary. Thus, these results are abstracted from the FDA Summary, which reports denominators for most results.

The overall success rate was 63.5% (94/148) for the ProDisc, and 45.1% (32/71) for fusion (p=0.0053). Using FDA alternate criteria for success (15-point change in ODI, and alternative radiographic success criteria) the overall success rate was 53.4% for the ProDisc and 40.8% for fusion (p=0.044). Note that the rates are calculated on fewer patients than the total randomized. Eleven percent of (9/80) fusion patients and 7.5% (12/160) of ProDisc patients have been excluded from the results, presumably due to missing data. The FDA Summary states that "the analysis population…consisted of all randomized subjects who completed all evaluations at the 24-month time point, regardless of when the 24-month measurement occurred." Separate components of the composite outcome are also reported using various denominators. The losses to follow-up are difficult to reconcile from information provided in a table of patient accountability.

In examining the individual components of the success rate, patients receiving the ProDisc had greater success in achieving improvement in the ODI (77.2% versus 64.8%, p=0.039). Contradicting this result somewhat is the finding that the mean ODI scores between the two groups were not statistically significant, nor were the VAS pain scores. The rate of neurologic success was higher in the ProDisc patients (91.2% versus 81.4%, p=0.0341).

Of concern in the reporting is whether device failures are included in the tabulation of overall success rates. During the trial, there were 2 fusion device failures and 6 ProDisc device failures. Although it is unclear whether these device failures were appropriately accounted for in the tabulation of overall success, it appears that they were definitely not included in the tabulation of the individual components. Several of the device failures involved movement of the disc and required reoperation, but the prepublication draft reports that the radiographic outcomes were reported on "…patients reaching 24 month follow-up without re-operation…."

Selected adverse events abstracted from the FDA Summary are shown in Table 9. Any type of adverse event was experienced by 87.5% of fusion patients and 84% of ProDisc patients. Twenty percent of fusion patients experienced a device-related adverse event and 17.9% of ProDisc patients experienced a device related event. Device failures, defined as those that required reoperation or revision occurred in 6 ProDisc patients and 2 fusion patients. In 3 of the patients, the polyethylene core dislodged. In one of these, it was thought that the core was not locked correctly at the time of surgery. In one case, the entire implant migrated. In another case, the polyethylene core was inserted incorrectly and required reinsertion. The last device failure was a reoperation for supplemental fixation due to continuing pain. The 2 fusion patients who failed required reoperation for unresolved pain. Two additional fusion patients who were not considered device failures had routine removal of hardware.

Summary of ProDisc

In sum, the evidence supporting the effectiveness of the ProDisc 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.

Noninferiority p-values were not presented in either the FDA Summary or the published trial, but traditional p-values that are less than 0.05 will meet noninferiority criteria. However, the same potential problems regarding the interpretation of a noninferiority trial apply here. Precise knowledge of the effectiveness of fusion is lacking. Lower-than-expected success rates (63.5% achieved versus 85% planned) make it uncertain whether meeting noninferiority criteria is consistent with ProDisc superiority to conservative management. The higher success rates achieved in both groups in improvement in Oswestry scores than in the Charité trial could in part be due to the less-stringent standard for improvement. When looking at the mean reduction in ODI scores, the two devices appear to produce similar mean reductions (Charité trial 48.5% reduction, ProDisc trial 45.6% reduction [estimated based on baseline ODI of 63.4 and final ODI of 34.5]).

If viewed from the perspective of a traditional superiority trial, straightforward interpretation of the p-values presented are hampered by lack of intent-to-treat analysis and uncertain handling of missing values and device failures.

Summary of Application of the Technology Evaluation Criteria

Based on the available evidence, the Blue Cross and Blue Shield Medical Advisory Panel made the following judgments about whether the artificial lumbar disc for treatment of degenerative disc disease meets the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) criteria.

1. The technology must have final approval from the appropriate governmental regulatory bodies.

In October 2004, the U.S. Food and Drug Administration (FDA) granted premarket application (PMA) approval for the Charité Artificial Disc, stating that the device is indicated for spinal arthroplasty in skeletally mature patients with DDD at one level from L4-S1. DDD is defined as discogenic back pain with degeneration of the disc confirmed by patient history and radiographic studies. As a condition of approval, the manufacturer has agreed to conduct a postapproval study, using a maximum of 366 patients (201 randomized investigational subjects, 67 training investigational subjects, and 98 control subjects). Postapproval study patients will be evaluated for a period of 5 years post-implantation.

The FDA granted approval to the ProDisc in August 2006 and is indicated for spinal arthroplasty in skeletally mature patients with DDD at one level from L3-S1. Patients should have no more than Grade 1 spondylolisthesis at the involved level. Patients receiving the ProDisc should have failed at least 6 months of conservative treatment prior to implantation. The manufacturer agreed to conduct a postapproval study to obtain 5-year follow-up data from all subjects in the clinical study. The study will utilize the same endpoints as the IDE clinical study, and also evaluate adjacent segment degeneration and correlation of range-of-motion with ODI and VAS scores.

2. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.

Current evidence supporting the effectiveness of artificial lumbar disc is insufficient. Case series evidence is inadequate to establish efficacy. The randomized trials of the Charité disc and ProDisc are suspect as valid noninferiority trials and do not prove superiority.

3. The technology must improve the net health outcome; and
4. The technology must be as beneficial as any established alternatives.

The evidence is insufficient to determine whether the use of artificial lumbar discs improves the net health outcome or whether they are as beneficial as any established alternatives.

5. The improvement must be attainable outside the investigational settings.

Whether the use of artificial lumbar discs improves health outcomes has not been established in the investigational settings.

Therefore, the use of artificial lumbar discs for degenerative disc disease does not meet the TEC criteria.

References

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

Bertagnoli R, Yue JJ, Shah RV et al. (2006a). 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,* 30:2192-9.

Bertagnoli R, Yue JJ, Nanieva R et al. (2006b). 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,* 4:85-90.

Blumenthal S, McAfee PC, Guyer RD et al. (2005). A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITÉ artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. *Spine,* 30:1565-75; Erratum in: *Spine* 2005, 30:2356.

Cherkin DC, Sherman KJ, Deyo RA et al. (2003). A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. *Ann Intern Med,* 138:898-906.

Christensen FB, Hansen ES, Eiskjaer SP et al. (2002). Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. *Spine,* 27:2674-83.

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

Cinotti G, David T, Postacchini F. (1996). Results of disc prosthesis after a minimum follow-up period of 2 years. *Spine,* 21:995-1000.

de Kleuver M, Oner FC, Jacobs WC. (2003). Total disc replacement for chronic low back pain: background and a systematic review of the literature. *Eur Spine J,* 12:108-16.

David T. (1993). Lumbar disc prosthesis: surgical technique, indications and clinical results in 22 patients with a minimum of 12 months follow-up. *Eur Spine J,* 1:254-259.

Deyo RA, Nachemson A, Mirza SK. (2004). Spinal-fusion surgery - the case for restraint. *N Engl J Med,* 350:722-6.

Deyo RA, Tsui-Wu YJ. (1987). Descriptive epidemiology of low-back pain and its related medical care in the United States. *Spine,* 12:264-268.

Fritzell P, Hagg O, Wessberg P et al. (2001). Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine,* 26:2521-32.

Geisler FH, Blumenthal SL, Guyer RD et al. (2004). Neurological complications of lumbar artificial disc replacement and comparison of clinical results with those related to lumbar arthrodesis in the literature: results of a multicenter, prospective, randomized investigational device exemption study of Charité intervertebral disc. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. *J Neurosurg Spine,* 1:143-54.

Gibson JN, Grant IC, Waddell G. (1999). The Cochrane review of surgery for lumbar disc prolapse and degenerative lumbar spondylosis. *Spine,* 24:1820-32.

Griffith SL, Shelokov AP, Buttner-Janz K et al. (1994). A multicenter retrospective study of the clinical results of the LINK SB Charité intervertebral prosthesis. The initial European experience. *Spine,* 19:1842-9.

Grigg P, Schaible HG, Schmidt RF. (1986). Mechanical sensitivity of group III and IV afferents from posterior articular nerve in normal and inflamed cat knee. *J Neurophysiol,* 55:635-643.

Haldeman S. (1999). Low back pain: current physiologic concepts. *Neurol Clin,* 17(1):1-15.

Hirsch C, Ingelmark BE, Miller M. (1963). The anatomical basis for low back pain. *Acta Orthop Scand,* 33:1-17.

Huang RC, Sandhu HS. (2004). The current status of lumbar total disc replacement. *Orthop Clin North Am,* 35(1):33-42.

Ivar Brox J, Sorensen R, Friis A et al. (2003). Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine,* 28:1913-21.

Kaul S, Diamond GA. (2006). Good enough: a primer on the analysis and interpretation of noninferiority trials. *Ann Intern Med,* 145:62-9.

Kuslich DS, Ahern JW. (1994). What tissues are responsible for low back pain and sciatica: an investigation of tissue sensitivity in humans. In: *Proceedings of the Orthopedic Research Society*, 40th Annual Meeting, New Orleans, LA, p. 110.

Lemaire JP, Skalli W, Lavaste F et al. (1997). Intervertebral disc prosthesis. Results and prospects for the year 2000. *Clin Orthop,* (337):64-76.

Madan SS, Boeree NR. (2003). Comparison of instrumented anterior interbody fusion with instrumented circumferential lumbar fusion. *Eur Spine J,* 12:567-75.

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

Meade T, Brown W, Mellows S et al. (1986). Comparison of chiropractic and outpatient management of low back pain: a feasibility study. *J Epidemiol Commun Health,* 40:12-17.

Pope MH, DeVocht JW. (1999). The clinical relevance of biomechanics. *Neurol Clin,* 17:17-41.

Roland M, Fairbank J. (2000). The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine,* 25:3115-24.

Schroven I, Dorofey D. (2006). Intervertebral prosthesis versus anterior lumbar interbody fusion: one-year results of a prospective non-randomised study. *Acta Orthop Belg,* 72:83-6.

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

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

Smyth MJ, Wright V. (1958). Sciatica and the intervertebral disc. *J Bone Joint Surg Am,* 40:1401-1418.

Sott AH, Harrison DJ. (2000). Increasing age does not affect good outcome after lumbar disc replacement. *Int Orthop*, 24:50-3.

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

Swenson R. (1999). Differential diagnosis: a reasonable clinical approach. *Neurol Clin,* 17:43-63.

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

Turner JA, Ersek M, Herron L et al. (1992). Patient outcomes after lumbar spinal fusions. *JAMA,* 268:907-11.

U.S. Food and Drug Administration (FDA) Center for Devices and Radiological Health. (2004a). DePuy Spine Charité™ Artificial Disc (P040006) clinical review. Available online at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/4049b1_03_Clinical%20Review%20Memo%20MAY.pdf. Last accessed January 2005.

U.S. Food and Drug Administration (FDA) Center for Devices and Radiological Health. (2004b). In-depth statistical review for expedited PMA (P040006) Charité™ Artificial Disc, DePuy Spine, Inc. (dated Feb. 23, 2004). Available online at http://www.fda.gov/ohrms/dockets/ac/04/briefing/4049b1_04_Statistical%20Review%20Memo%20JCC.pdf. Last accessed January 2005.

U.S. Food and Drug Administration (FDA) Center for Devices and Radiological Health. (2004). Premarket application approval letter from Director, Office of Device Evaluation (Tillman DB) to Vice President, Clinical and Regulatory Affairs, DePuy Spine, Inc. (Christianson W), dated October 26, 2004. Available online at: http://www.fda.gov/cdrh/pdf4/p040006a.pdf. Last accessed February 2005.

U.S. Food and Drug Administration (FDA) Center for Devices and Radiological Health. (2005). Summary of safety and effectiveness: Charité artificial disc. Available online at http://www.fda.gov/ohrms/dockets/dockets/05m0092/05m-0092-aav0001-03-SSED-vol1.pdf.

U.S. Food and Drug Administration (FDA) Center for Devices and Radiological Health. (2006). Summary of safety and effectiveness: ProDisc-L total disc replacement. Available online at http://www.fda.gov/cdrh/pdf5/p050010b.pdf.

van Ooij A, Kurtz SM, Stessels F et al. (2007). Polyethylene wear debris and long-term clinical failure of the Charité disc prosthesis: a study of 4 patients. *Spine,* 32:223-9.

Videbaek TS, Christensen FB, Soegaard R et al. (2006). Circumferential fusion improves outcome in comparison with instrumented posterolateral fusion: long-term results of a randomized clinical trial. *Spine,* 31:2875-80.

Weinstein J, Claverie W, Gibson S. (1988). The pain of discography. *Spine,* 13:1344-1348.

Yoshizawa H, O'Brien JP, Smith WT et al. (1980). The neuropathology of intervertebral discs removed for low back pain. *J Pathol,* 132:95-104.

Zeegers WS, Bohnen LM, Laaper M et al. (1999). Artificial disc replacement with the modular type SB Charité III: 2-year results in 50 prospectively studied patients. *Eur Spine J,* 8:210-7.

Zigler J, Delamarter R, Spivak JM et al. (2007). 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,* 32(11):1155-62.

Technology
Evaluation
Center

Blue Cross and
Blue Shield Association
225 North Michigan Avenue
Chicago, Illinois 60601-7680
www.bcbs.com/tec