

Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITÉ artificial disc versus lumbar fusion: Five-year follow-up

Richard D. Guyer, MD^{a,*}, Paul C. McAfee, MD^b, Robert J. Banco, MD^c, Fabian D. Bitan, MD^d, Andrew Cappuccino, MD^e, Fred H. Geisler, MD PhD^f, Stephen H. Hochschuler, MD^a, Richard T. Holt, MD^g, Louis G. Jenis, MD^c, Mohamed E. Majd, MD^g, John J. Regan, MD^h, Scott G. Tromanhauser, MD^c, Douglas C. Wong, MDⁱ, Scott L. Blumenthal, MD^a

^aTexas Back Institute, 6020 West Parker Road, Suite 200, Plano, TX 75093, USA

^bSpine and Scoliosis Center, 7505 Osler Drive, Suite 104, Towson, MD 21204, USA

^cBoston Spine Group, 125 Parker Hill Avenue, Boston, MA 02120, USA

^dLennox Hill Hospital, 130 East 77th Street, 7th Floor, New York, NY 10003, USA

^eBuffalo Spine Surgery, 46 Davison Street, Lockport, NY 14094, USA

^fIllinois Neuro-Spine Center, 2020 Ogden Avenue, Suite 335, Aurora, IL 60504, USA

^gSpine Surgery, PSC, 210 East Gray Street, Suite 601, Louisville, KY 40202, USA

^hG&P Orthopedic Associates, 120 South Spalding Drive, Suite 400, Beverly Hills, CA 90212, USA

ⁱPanorama Orthopedics, 660 Golden Ridge Road, Suite 250, Golden, CO 80401, USA

Received 9 April 2008; accepted 5 August 2008

Abstract

BACKGROUND CONTEXT: The CHARITÉ artificial disc, a lumbar spinal arthroplasty device, was approved by the United States Food and Drug Administration in 2004 based on two-year safety and effectiveness data from a multicenter, prospective, randomized investigational device exemption (IDE) study. No long-term, randomized, prospective study on the CHARITÉ disc or any other artificial disc has been published to date.

PURPOSE: The purpose of this study was to compare the safety and effectiveness at the five-year follow-up time point of lumbar total disc replacement using the CHARITÉ artificial disc (DePuy Spine, Raynham, MA) with that of anterior lumbar interbody fusion (ALIF) with BAK cages and iliac crest autograft, for the treatment of single-level degenerative disc disease from L4 to S1, unresponsive to nonoperative treatment.

STUDY DESIGN/SETTING: Randomized controlled trial—five-year follow-up.

PATIENT SAMPLE: Ninety CHARITÉ patients and 43 BAK patients.

FDA device/drug status: approved for this indication (CHARITÉ disc).

Author disclosures: RDG (research support and royalties from DePuy Spine; stockholder at K2M, Spinal Motion; consultant for DePuy Spine, Spinal Motion; member of scientific advisory board at K2M; clinical staff or training receives educational support from DePuy spine, Abbott Spine, Synthes Spine, Medtronic, Neuro); PCM (research support and royalties; stockholder; consultant and speaker; received support for staff and materials, discretionary funds, trips/travel); RJB (consultant and speaker); FDB (consultant and speaker; supported for trips/travel); AC (receives royalties and stockholder at JJ DePuy, NuVasive, Pioneer); FHG (royalties from Aescular, DePuy Spine; stockholder at Disc Motion, Impliant Surgitech, Spinal Integration; consultant for DePuy Spine, Cervitech, Abbott Spine, K2K and speaker for DePuy Spine; in scientific advisory board at Disc Motion, K2M; received support for staff and materials from DePuy Spine, Cervitech, Spinal Motion, Implic); SHH (stockholder, IST,

Spinewave Replication Medical, K2M, Pioneer Surgical Spinal Motion, Disc Motion, Anvlex, Archus, Anson; consultant for Orthotic, Abott Spine, DePuy Spine, Alphatec, Endios; member of board of directors (Alphatec, SpineMark), scientific advisory board (Pioneer), faculty for fellowship receiving educational/research support from Abbott Spine, DePuy Spine, Synthes, Spine & Medtronic Neuro); LGJ (consultant for DePuy Spine); JJR (royalties from Johnson & Johnson); SGT (consultant for DePuy, Medtronic, Stryker, Zimmer; received research support from Impliant, Stryker, Zimmer); DCW (royalties from Biomet/EBE; stockholder at Synthes; consultant for DePuy, speaker for Kyphen); SLB (receives royalties and consultant for DePuy Spine).

* Corresponding author. Texas Back Institute, 6020 West Parker Road, Suite 200, Plano, TX 75093, USA. Tel.: (972) 608-5114; fax: (972) 608-5020.

E-mail address: rguyer@texasback.com (R.D. Guyer)

OUTCOME MEASURES: *Self-reported measures:* visual analog scale (VAS); validated Oswestry disability index (ODI version 1.0); Short-Form 36 Questionnaire, and patient satisfaction. *Physiologic measures:* radiographic range of motion, disc height, and segmental translation. *Functional measures:* work status.

METHODS: Of the 375 subjects enrolled in the CHARITÉ IDE trial, 277 were eligible for the five-year study and 160 patients thereof completed the five-year follow-up. The completers included 133 randomized patients. Overall success was defined as improvement ≥ 15 pts in ODI vs. baseline, no device failure, absence of major complications, and maintenance or improvement of neurological status. Additional clinical outcomes included an ODI questionnaire as well as VAS, SF-36, and patient satisfaction surveys. Work status was tracked for all patients. Safety assessments included occurrence and severity of adverse events and device failures. Radiographic analyses such as index- and adjacent-level range of motion, segmental translation, disc height, and longitudinal ossification were also carried out.

RESULTS: Overall success was 57.8% in the CHARITÉ group vs. 51.2% in the BAK group (Blackwelder's test: $p=0.0359$, $\Delta=0.10$). In addition, mean changes from baseline for ODI (CHARITÉ: -24.0 pts vs. BAK: -27.5 pts), VAS pain scores (CHARITÉ: -38.7 vs. BAK: -40.0), and SF-36 questionnaires (SF-36 Physical Component Scores [PCS]: CHARITÉ: 12.6 pts vs. BAK: 12.3 pts) were similar across groups. In patient satisfaction surveys, 78% of CHARITÉ patients were satisfied vs. 72% of BAK patients. A total of 65.6% patients in the CHARITÉ group vs. 46.5% patients in the BAK group were employed full-time. This difference was statistically significant ($p=0.0403$). Long-term disability was recorded for 8.0% of CHARITÉ patients and 20.9% of BAK patients, a difference that was also statistically significant ($p=0.0441$). Additional index-level surgery was performed in 7.7% of CHARITÉ patients and 16.3% of BAK patients.

Radiographic findings included operative and adjacent-level range of motion (ROM), intervertebral disc height and segmental translation. At the five-year follow-up, the mean ROM at the index level was 6.0° for CHARITÉ patients and 1.0° for BAK patients. Changes in disc height were also similar for both CHARITÉ and BAK patients (0.7 mm for both groups, $p=0.9827$). Segmental translation was 0.4 and 0.8 mm in patients implanted with CHARITÉ at L4–L5 vs. L5–S1, respectively, and 0.1 mm in BAK patients.

CONCLUSIONS: The results of this five-year, prospective, randomized multicenter study are consistent with the two-year reports of noninferiority of CHARITÉ artificial disc vs. ALIF with BAK and iliac crest autograft. No statistical differences were found in clinical outcomes between groups. In addition, CHARITÉ patients reached a statistically greater rate of part- and full-time employment and a statistically lower rate of long-term disability, compared with BAK patients. Radiographically, the ROMs at index- and adjacent levels were not statistically different from those observed at two-years postsurgery. © 2009 Elsevier Inc. All rights reserved.

Keywords:

Arthroplasty; Arthrodesis; 5-year follow-up; Randomized controlled trial

Introduction

Spinal arthroplasty has been studied for a number of years as an alternative to spinal arthrodesis in the surgical treatment of degenerative disc disease (DDD). The first arthroplasty device, the SB CHARITÉ I, was implanted as early as in 1984 [1]. Since that time, this prosthesis underwent two major design changes [2] and, in its final form, was thoroughly tested for biomechanical integrity [3] and in vivo effectiveness [4] before clinical evaluation.

The first clinical experience with the CHARITÉ III, also known as the CHARITÉ artificial disc, was reported in 1994 by Griffith et al. [5]. This retrospective study analyzed charts from 93 patients and reported good overall symptom relief with respect to back and leg pain at a one-year follow-up time point. Soon thereafter, another retrospective review including a 45-patient cohort described outcomes at 3.2 years postoperatively [6]. In this study, clinical improvements

were again satisfactory; however, authors highlighted the need for prospective randomized studies with longer-term follow-up periods to assess the validity of the clinical results obtained so far and to compare arthroplasty with fusion.

The recent publications of several mid- and long-term arthroplasty studies represent the closest attempts so far at addressing this shortfall. Lemaire et al. first reported 10-year follow-up results in 100 patients [7]. Overall, the authors reported excellent or good clinical outcomes in 90% cases. Adjacent-level degeneration was observed in only 2% of cases, whereas the rate of facet joint degeneration remained relatively minimal (11%).

These promising results were challenged by Putzier et al. in 2006 in a second, long-term study, reporting 17-year follow-up data with CHARITÉ [8]. In this report, 60% spontaneous ankylosis was observed, prompting authors to conclude that the long-term efficacy of total disc replacement still needed to be examined. Although this

EVIDENCE & METHODS

Context

The Charite artificial disc has been used in North America for a limited time, with continued questions regarding the relative advantages and durability of the prosthesis compared to spinal fusion.

Contribution

The authors present the five-year follow-up on “available” patients from a multicenter investigational device exemption (IDE) trial, comparing the Charite artificial disc to the largely abandoned BAK fusion device. Only some subjects (58%) were contacted at five years, and some sites in the original trial declined to participate at all. In the select patients contacted, the five-year outcomes were very similar between the two Charite and BAK patients. Approximately 55% of the reporting patients met the “success” standard (which included rather minimal improvements in function.)

Implications

Caution is advised. By reporting on only a partial cohort sample, not randomly selected, this study design introduces a risk of surveillance and reporting bias. The comparison to the BAK device, with its acknowledged limitations, certainly suggests a methodological bias against fusion compared to modern techniques. The outcomes in either group are somewhat disappointing. Even in these highly selected patients, it is unclear these rates are acceptable. Finally the only device failures in the two- to five-year period were in the Charite group. Fortunately, although there was a trend to more re-operations in the Charite group in that time frame, that rate (2%) was small in the select patients followed.

—The Editors

study presented significant limitations, such as use of undersized prototype devices, suboptimal device placement, ill-defined patient selection, and incorrect statistical calculations [1,2], it further reinforced the need for a clinically relevant, long-term safety and efficacy study, comparing the CHARITÉ device with an established fusion procedure. In absence thereof, however, David published a third long-term study, presenting a 10-year data on 106 patients with one-level DDD [9]. Excellent or good clinical outcome was obtained in 82.1% of patients, with an 89.6% rate of return to work. The authors thus concluded that disc replacement with CHARITÉ was a viable alternative to arthrodesis for the treatment of lumbar DDD.

Concurrent to Lemaire et al.’s publication were two disclosures of the CHARITÉ investigational device exemption (IDE) trial designed to evaluate the safety and effectiveness of the CHARITÉ artificial disc vs. anterior lumbar interbody fusion (ALIF) with BAK and autograft [10,11]. These reports demonstrated the noninferiority of the CHARITÉ artificial disc to the BAK fusion at the two-year time point. The United States Food and Drug Administration (FDA) granted approval to the CHARITÉ artificial disc on the basis of these results and requested continuation of the study to a five-year follow-up time point.

The results that follow describe the five-year clinical and radiographic outcomes for the CHARITÉ IDE multicenter, randomized controlled trial (RCT). This study represents the first long-term multicenter RCT comparing arthroplasty with arthrodesis.

Materials and method

Study design

Between May 2000 and April 2002, 375 patients had surgery in a prospective, randomized, nonblinded, FDA-approved study at 14 investigational sites across the United States. Institutional Review Board approval, patient enrollment, inclusion and exclusion criteria were previously described by Blumenthal et al. [10]. Patients were randomly assigned in a 2:1 ratio to one of two groups. The investigational group was implanted with the CHARITÉ Artificial Disc, whereas the control group was treated by ALIF with BAK threaded fusion cages (Zimmer Spine, Minneapolis, MN) packed with iliac crest autograft. Each site was given the opportunity to treat up to five nonrandomized CHARITÉ cases.

Of the 14 initial sites, 6 declined participation in the 5-year continuation study, reducing the total number of eligible patients by 90, 64 randomized and 26 training cases. Four deaths and 4 device removals reduced the overall eligible patient population to 277 patients (233 randomized and 44 nonrandomized). Loss to follow-up included 117 patients, of whom 11 declined continuation in the study and 10 had early discontinuation. A total of 160 patients completed the five-year study, including 27 nonrandomized training cases and 133 randomized cases (90 CHARITÉ and 43 BAK patients). Thus, the follow-up rates reached 57% of eligible randomized patient population and 44% of the total IDE patient cohort. Results from the nonrandomized training cases were not included herein—this report focuses exclusively on results from randomized cases. Protocols for this five-year follow-up study were defined prospectively, before any of the patients were seen. This study is therefore a prospective study.

Patient demographics

Before surgery, most of the patients enrolled in the study met the inclusion and exclusion criteria; however,

osteopenia (an exclusion criterion) was seen in five patients from the CHARITÉ group and one patient from the BAK group. There was no significant difference between the two groups with respect to gender, age, race, height, body mass index, incidence of prior spinal surgery, activity level before the onset of symptoms, activity level at the time of enrollment, or preoperative working status (Table 1).

Clinical outcome measurements

Clinical evaluations were completed before surgery and at 6 weeks and 3, 6, 12, 24, and 60 months postsurgery. They included a visual analog scale (VAS) for pain, the Oswestry disability index (ODI), neurological status and short-form 36 (SF-36) health survey questionnaire, as well as work status evaluations. The VAS used here ranged from 0 to 100. The ODI tool consisted of the standard ten-item ordinal scale with six possible responses for each item and a total score ranging from 0 to 100 (Version 1.0). For both, VAS and ODI, lower scores were indicative of less pain/disability. The SF-36 questionnaire was similar to the previously validated and discussed version [12]. PCS results are shown herein. Clinical success as defined by the FDA (nonvalidated clinical scale) at the five-year time point was achieved when all the following four criteria were met: 1) ≥ 15 pts improvement in ODI; 2) no device failure; 3) no major complication, and 4) no neurological change. Device failure included any case of unplanned return to the operating room for index-level surgery, that is, supplemental fixation, decompression or reoperation, and thus addressed device-related safety concerns. Major complications were defined as major vessel injury resulting in >1500 cc blood loss, neurological damage, or nerve root injury. No neurological change was defined as lack of neurological deterioration compared with preoperative status, at any point of time.

Additional surgery for adjacent-level degeneration was recorded. In addition, the number of patients under pain management for adjacent-level disc disease was also analyzed and compared between groups.

Surgical technique

All patients underwent surgical treatment through an open anterior retroperitoneal approach. The surgical technique for implantation of the CHARITÉ artificial disc has been described previously in great detail by Geisler [13] and more recently, in the two-year follow-up release of this study [10]. Briefly, patients were placed in the supine position. Fluoroscopy was used to identify the approach angle and location of the disc space. The approach was performed in standard fashion with the assistance of an access surgeon in most of the cases. A complete discectomy was performed using standard anterior lumbar surgical instruments. Special instruments were used to assess proper footprint sizing, lordotic angle, core height, and placement of the

Table 1
Patient demographics

Variable	CHARITÉ group (n=90)	BAK group (n=43)	p Value
Gender, n (%)			
Male	47 (52%)	24 (56%)	0.7145
Female	43 (48%)	19 (44%)	
Age (y)			
Mean (SD)	40.0 (8.58)	38.8 (8.69)	0.4673
Median	40.5	39.0	
Range	19–60	25–55	
>45 y, n (%)	23 (26%)	9 (21%)	0.6666
≤ 45 y, n (%)	67 (74%)	34 (79%)	
Race, n (%)			
Caucasian	85 (94)	40 (93)	0.8463
African American	3 (3)	2 (5)	
Other	2 (2)	1 (2)	
Height (cm)			
Mean (SD)	173.7 (9.46)	173.4 (8.94)	0.8717
Median	175.3	172.7	
Range	150–193	155–191	
Weight (kg)			
Mean (SD)	80.0 (16.08)	81.6 (15.87)	0.5968
Median	81.0	79.8	
Range	51–120	54–118	
Body mass index			
Mean (SD)	26.4 (4.13)	27.1 (4.79)	0.3704
Median	26.3	26.9	
Range	17–37	19–40	
Previous spinal surgeries, n (%)			
Yes	37 (41%)	12 (28%)	0.1792
No	53 (59%)	31 (72%)	
Normal activity level, n (%) (before experiencing back pain)			
Active	84 (93%)	37 (86%)	0.2237
Moderate	6 (7%)	5 (12%)	
Light	0	1 (2%)	
Minimal	0	0	
Activity level at enrollment, n (%)			
Active	5 (6%)	0	0.2963
Moderate	14 (16%)	4 (9%)	
Light	22 (24%)	10 (23%)	
Minimal	49 (54%)	29 (67%)	
Presurgery work status, n (% working)			
Full-time	46 (51%)	21 (49%)	0.8728
Part-time	6 (7%)	3 (7%)	
Short-term disability	2 (2%)	3 (7%)	
Long-term disability	9 (10%)	4 (9%)	
Not employed	10 (11%)	4 (9%)	
Retired	2 (2%)	0	
Other	15 (17%)	8 (9%)	

Fisher exact test used to test categorical variables and *t* test used to test means.

prosthesis within the disc space under fluoroscopy. The prosthesis end plates were then inserted into the disc space in a trajectory parallel to the vertebral end plates as determined from lateral fluoroscopy. Final positioning was assessed with fluoroscopy.

Disc space preparation was identical for patients in both groups. The anterior longitudinal ligament and anterior annulus fibrosis were resected. A complete discectomy was performed with preservation of the peripheral annulus fibrosis to provide ligamentotaxis. Care was taken not to violate the bony end plates, and in some cases, posterior osteophytes were carefully removed to allow satisfactory placement of the prosthesis. The posterior longitudinal ligament was stretched to facilitate restoration of normal disc space height. In the BAK control group, instruments specific to the BAK cages were used for sizing of the disc space. Corticocancellous autograft was harvested from the iliac crest and packed into two BAK cages. The cages were placed in the disc space according to the recommended surgical technique. In both groups, the wound was closed in standard fashion and patients were taken to the recovery room.

Radiographic analyses

The detail of the radiographic scanning technique was previously discussed for the full IDE patient population by McAfee et al. [11]. Briefly, all patients were X-rayed before surgery and later at all time points (6 weeks and 3, 6, 12, 24, and 60 months). Radiographs included anteroposterior (AP), lateral and full flexion, and extension views. The radiographic scanning technique was based on that reported by Kuslich et al. [14]. All radiographs were scanned, digitized, and analyzed by a software program designed to measure differences between flexion/extension angles and translations of the operative motion segment using validated, computer-assisted methods, with mean error less than 0.5° and confidence interval range of 0.3° , as defined in cadaver studies [15] (QMA, Medical Metrics, Houston, TX). All radiographs were analyzed by the same system, to avoid inter- and intraobserver variability. Measurements on all radiographs included flexion/extension ROM at operated and adjacent levels, segmental translation as well as longitudinal disc height. Segmental translation, which can be described as an AP movement of the two spinal vertebrae in relationship to each other, was analyzed using the motion of the posterior inferior corner of the superior vertebra in a direction defined by the superior end plate of the inferior vertebra. Motion was measured parallel to the superior end plate of the inferior vertebra.

Longitudinal ossification was also evaluated on all X-rays by board-certified radiologists at Medical Metrics Inc. For CHARITÉ cases, a 6-point scale was setup from a combined classification system adapted from Nathan [16] and McAfee et al. [17]. The grades were defined as follows: 0: no evidence of osteophyte formation or heterotopic ossification; 1: osteophytes/heterotopic ossification appearing as isolated points of initial hyperostosis or islands of bone in soft tissue; 2: bone occurring within the disc space defined by the planes formed by the two adjacent end plates, but bony protrusions projecting more or less

horizontally from the vertebral bodies; 3: bone occurring between the two planes without bridge; 4: apparent continuous connection of bridging bone between end plates; 5: indeterminate, for example, due to poor film quality. To determine the clinical relevance of the longitudinal ossification, segments rated 3 or 4 were further correlated to overall segment motion. Longitudinal ossification was clinically relevant if it prevented motion. Motion was defined as any movement greater than a specific cut-off value. Two different cut-off values were used: 1) a 5° cut-off value, as per FDA guidelines [18]; and 2) a 3° cut-off value, as per FDA assessment of motion in the ProDisc L summary of safety and effectiveness [19].

Statistical methods

Data were analyzed using the SAS v8.2 statistical software package (SAS Institute, Cary, NC). For categorical variables, p values were generated using the Fisher exact test. A *t* test was used to test means. The primary effectiveness evaluation for the study was an assessment of the equivalence of the two treatments in terms of the proportion of patients in each treatment group classified as a responder.

The success status of patients was summarized by treatment group using counts and percentages. The Blackwelder's test was used to test for treatment equivalence based on the assumption that a difference of 0.10 was clinically significant. Covariate analyses were done to assess the impact of various factors (age, baseline ODI, gender, operative level, use of hormone replacement therapy, use of pain medication at any time, body mass index, and baseline activity level).

To further understand the impact of missing data on the results of this study, two statistical analyses were performed. The first analysis was conducted comparing the five-year completers with all the patients who were lost to follow-up. Baseline and two-year clinical outcomes (VAS and ODI) of the five-year completers were statistically compared with those of patients lost to follow-up. A secondary analysis was conducted to determine whether a difference existed in baseline and two-year clinical outcome of patients at the nine participating sites vs. the six nonparticipating sites. A Fisher exact or chi-square test was used to test categorical variables and a *t* test was used to compare univariate summaries.

To confirm overall clinical success, results obtained using the Blackwelder's methodology and two additional statistical analyses were conducted. First, a Last Observation Carried Forward (LOCF) analysis was done using all available data, up to the latest follow-up time point for all patients (including cases of discontinuation, early removals, and deaths). Secondly, a Bayesian statistical analysis was performed. For this purpose, a predictive model was built based on the existing data of all available patients at two- and five-year follow-up time periods.

Results

Results from patients who crossed over to a different treatment group were maintained in the intended treatment category, and all results are therefore presented on an “intent-to-treat” versus “as-treated” basis.

Surgery

There was no difference between the two groups with respect to operative time, blood loss, or level of implantation (Table 2). Mean operative time was 108.2 minutes in the CHARITÉ group and 122.1 minutes in the BAK group (p=0.1614). Mean estimated blood loss was 212.1 mL in the CHARITÉ group and 204.3 mL in the BAK group (p=0.8644). As shown with the p values, the operative time and blood loss were not statistically different between the two groups.

Clinical outcomes

Oswestry disability index

The ODI results were plotted as mean values at each time point for both the CHARITÉ and the BAK groups (Fig. 1). In addition, the percentage of patients with at least a 15-pt improvement in ODI score vs. baseline was tabulated per group and time point (Fig. 2). The cut-off value of 15-pt was specified by the FDA and a prior publication by Hägg et al., indicated that the minimal clinically important difference (MCID) in ODI was 10 pts, a value that should be exceeded for clinical decision-making [20].

ODI values significantly improved at all postoperative time points compared with baseline. The MCID for ODI was reached at three-months postoperatively in both the CHARITÉ and BAK groups. At that time point, the change in mean ODI scores was –21.5 pts in the CHARITÉ group

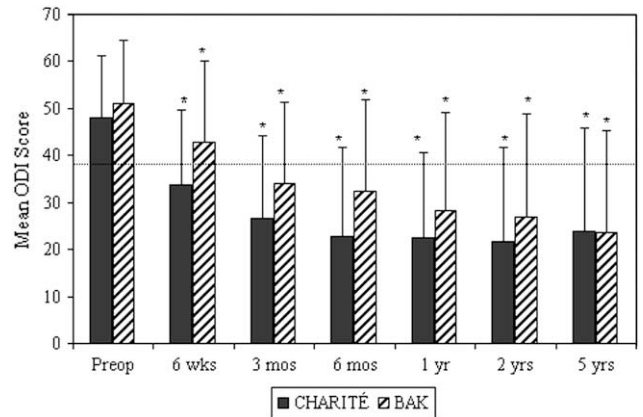


Fig. 1. Mean Oswestry scores for both groups and at all time points. Error bars represent the standard deviation. *Statistically significant difference from preoperative state within the same treatment group. The dotted line is placed at approximately 10-pt improvement from the CHARITÉ baseline and represents the MCID. Significant changes were observed as early as three-months postoperatively for both groups. Improvements remained constant between the two- and five-year time points.

and –17.0 pts in the BAK group. Mean ODI scores remained below 30 pts from one- to five-years postoperatively, in both groups. There was no statistical difference between the groups in terms of ODI scores, at the two- and five-year postoperative time points. Similarly, the percentage of patients with at least 15-pt improvement in ODI was equivalent at both the two- and five-year time points (Fig. 2).

VAS pain scores

VAS scores were plotted as mean values at each time point for both the CHARITÉ and the BAK groups (Fig. 3). The MCID value for VAS, as determined by Hägg et al., was estimated to be between 18 and 19 from preoperative scores [20].

VAS scores (range: 0–100) were significantly lower at all postoperative time points compared with baseline. The VAS

Table 2
Surgical procedure

Summary	CHARITÉ group (n=90)	BAK group (n=43)	p Value
Level fused or implanted			
L4/L5	26 (29%)	10 (23%)	0.5383
L5/S1	64 (71%)	33 (77%)	
Total surgery time (min)			
Mean (SD)	108.2 (45.99)	122.1 (66.29)	0.1614
Median	95.0	96.0	
Min, max	45.0, 250.0	55.0, 355.0	
Estimated blood loss (cc)			
Mean (SD)	212.1 (236.14)	204.3 (268.71)	0.8644
Median	100.0	120.0	
Min, max	25.0, 1500.0	20.0, 1600.0	
Hospital stay (day)			
Mean (SD)	3.7 (0.90)	4.3 (1.78)	0.0079
Median	4.0	4.0	
Min, max	1, 6	2, 12	

Fisher exact test used to test categorical variables and t test used to test means.

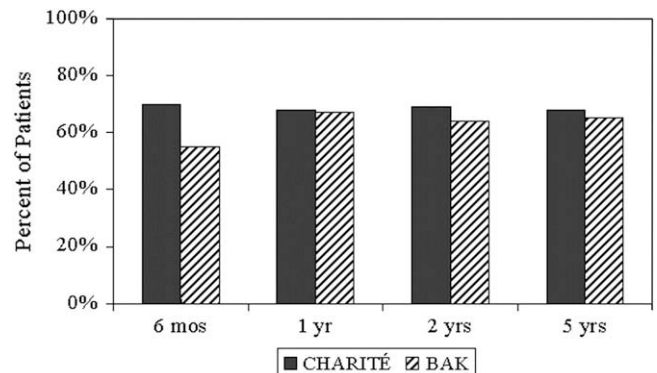


Fig. 2. Percentage of patients with at least a 15-pt improvement in ODI, for each group and at all time points. The percentage of patients with a 15-pt improvement exceeded 50% at six-months postoperatively and remained above the 50% threshold at both, the two- and five-year time points.

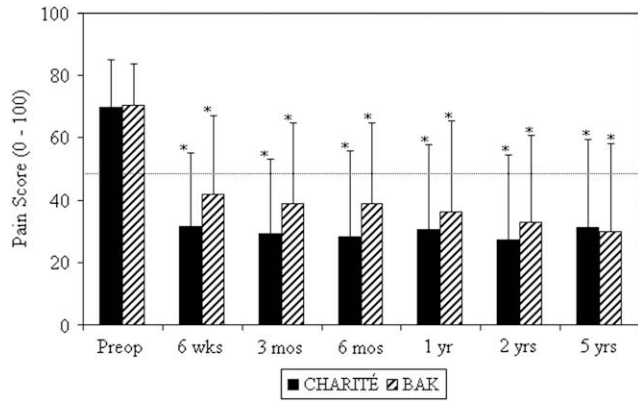


Fig. 3. Mean VAS scores for both groups and at all time points. Error bars represent the standard deviation. *Statistically significant difference from preoperative state within the same treatment group. The dotted line is placed at approximately 19-pt improvement from the CHARITÉ baseline and represents the MCID. VAS improvements exceeded MCID for both groups at the six-weeks postoperative time point and remained constant throughout the two- and five-year time points.

MCID was reached at the immediate postoperative time point in the CHARITÉ group (mean change in VAS scores: -23.1) and at six-weeks postsurgery in the BAK group. At that time point, the difference in mean VAS scores was -38.6 in the CHARITÉ group and -29.1 in the BAK group. The average VAS remained at approximately 30 from 2- to 5-years postsurgery, in both groups. There was no statistical difference between the groups in terms of VAS scores, at the two- and five-year postoperative time points.

Both VAS and ODI reached, and exceeded, MCID values early on postoperatively and maintained a clinically significant advantage over baseline throughout the 5-year follow-up time point.

SF-36 health questionnaires

The results of the SF-36 PCS were plotted as mean changes from baseline, at each time point for both the CHARITÉ and the BAK groups (Fig. 4). Although most

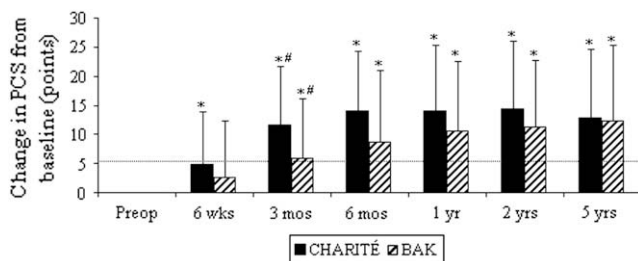


Fig. 4. Mean improvements in SF-36 PCS scores for all time points and both groups. Error bars represent the standard deviation. *Statistically significant difference from preoperative state within the same treatment group. #Statistically significant difference between treatment groups. The dotted line is placed at a 5.42-pt improvement from the CHARITÉ baseline and represents the MCID. Both groups exceeded 11 pts improvement at two- and five-years postoperatively. No statistical difference was observed between groups at any of the time points.

published spine surgeries provide an average of 8- to 16-pt improvement in SF-36 PCS, the MCID for SF-36 PCS scores has been established at a 5.42-pt improvement [21]. SF-36 PCS scores gradually improved at all postoperative time points compared with baseline. In the CHARITÉ group, increases in SF-36 PCS scores reached 14.2 pts at 2-year follow-up and 12.6 pts at 5-year follow-up. These results were similar to those observed in the BAK group (increases of 11.2 pts and 12.3 pts at two- and five-year follow-up, respectively). There was no statistical difference between the groups in terms of SF-36 PCS scores, at the two- and five-year postoperative time points.

Clinical success

As previously described, the following four criteria for clinical success are needed to be met to demonstrate individual clinical success: improvement ≥15 pts in ODI vs. baseline, no device failure requiring additional surgery, absence of major complications, and maintenance or improvement of neurological status. Table 3 outlines all the device failures observed throughout the five-year study. All device failures occurred in the BAK group between the 0 and 2-year time points (seven out of seven cases), whereas two out of a total of seven CHARITÉ device failures occurred between the 2- and 5-year time points.

Of the seven BAK failures, one case required a bilateral hemi-laminectomy with anterior discectomy and posterior fusion due to a non-union. The remaining six cases were treated with supplemental segmental instrumentation for posterior fixation due to pseudoarthroses (four out of six cases), facet joint arthrodesis (one out of six cases), and undefined persistent back pain (one out of six cases).

Of the seven CHARITÉ failures, one case required a reoperation without additional internal fixation (ie, right sided hemi-laminotomy, foraminotomy, and partial discectomy), whereas the remaining six cases were treated with supplemental segmental instrumentation for posterior fixation. The cause for these six supplemental fixations was as follows: symptomatic spondylolisthesis at L5 pars interarticularis, observed at six-month follow-up (one out of six

Table 3
Index-level surgery: device failures

	CHARITÉ group (n=90)	BAK group (n=43)	p Value
Summary			
0 to 2 ^a y	5 cases (5.5%) 5 supplemental fixations	7 cases (16.3%) 6 supplemental fixations 1 reoperation	
2 ^a to 5 y	2 cases (2.2%) 1 supplemental fixation 1 reoperation		
Total	7 cases (7.7%)	7 cases (16.3%)	0.1442

For p value: Fisher exact test was used to test categorical variables.
^aTwo-year window defined as <1020 days.

cases); device subsidence and subsequent low back pain (one out of six cases); facet degeneration (two out of six cases), and early postoperative implant displacement followed by back pain at the 12-month time point (two out of six cases). Both the cases of facet degeneration were diagnosed at the two-year follow-up time point.

The “overall clinical success,” a nonvalidated criterion defined by the FDA, was achieved in 57.8% patients in the CHARITÉ group and 51.2% patients in the BAK group at the five-year follow-up, compared with 65.2% and 60.6% of these same cohorts, respectively, at the two-year follow-up. The five-year overall success was further evaluated using a Blackwelder’s hypothesis (with $\Delta=0.10$, $p=0.0359$). This analysis provided a high degree of confidence that the CHARITÉ group was noninferior to the BAK group. The breakdown for the four separate clinical success factors is shown in Table 4.

Overall success was also calculated including all the four deaths and four early device removals as “failures.” Using this methodology, overall success was 54% for CHARITÉ and 50% for BAK (Blackwelder’s test with $\Delta=0.1$; $p=0.0670$).

Additional surgery for adjacent-level disease occurred in one CHARITÉ patient (1.1%) and two BAK patients (4.7%). In addition, nonsurgical treatment (eg, pain management) for DDD was provided to four CHARITÉ patients (4.4%) and six BAK patients (13.9%).

Work status

Work status was assessed before surgery, as shown in Table 1, and at the two- and five-year follow-ups. The exact

Table 4
Success rate for efficacy outcome by individual components

Summary	CHARITÉ group (n=90)	BAK group (n=43)	p Value
Fifteen-point improvement in Oswestry			
No	29 (32%)	15 (35%)	0.8443 ^a
Yes	61 (68%)	28 (65%)	
Device failures			
Failure	7 (8%)	7 (16%)	0.1442 ^a
Success	83 (92%)	36 (84%)	
Major complications			
Failure	0	0	
Success	90 (100%)	43 (100%)	
Neurological deterioration			
Failure	13 (14%)	7 (16%)	0.7989 ^a
Success	77 (86%)	36 (84%)	
Overall success	52 (57.8%)	22 (51.2%)	0.0359 ^b

Device failures were defined as requiring revision, reoperation, or removal. Major complications were defined as major vessel injury, neurological damage, nerve root injury, and death. Neurological deterioration included slight deterioration, significant deterioration, or mixed response at 60 months.

^aFor p values: Fisher exact test was used to test categorical variables.

^bBlackwelder’s test of noninferiority.

timing when an individual returned to work was not captured beyond these three time points. Work and disability status at the 5-year time point are presented in Fig. 5. Full-time employment was achieved by 65.6% patients in the CHARITÉ group vs. 46.5% patients in the BAK group. This difference was statistically significant ($p=0.0403$). In the CHARITÉ group, 8.0% patients were on long-term disability vs. 20.9% patients in the BAK group. This difference again was statistically significant ($p=0.0441$). No statistical significance was observed for all other parameters (part-time employment, short-term disability, not employed, retired, other).

Radiographic results

Range of motion

For CHARITÉ and BAK patients implanted at L4–L5, the ROM was evaluated at L4–L5 as well as L3–L4 and L5–S1. For CHARITÉ and BAK patients implanted at L5–S1, the ROM was evaluated at L5–S1 and L4–L5. All measurements were analyzed in terms of means and medians for all levels and treatment groups. Statistical analyses provided equivalent findings, where means or medians were used; therefore, to maintain consistency with prior publications, all analyses described below are based on means. Index-level range of motion (ROM) is shown in Fig. 6.

The mean preoperative ROM at L4–L5 for all patients (CHARITÉ and BAK) implanted at that level was 8.8°. There was no statistical difference in preoperative L4–L5 ROM between these CHARITÉ and BAK patients (ROM=8.7° for CHARITÉ patients and 9.2° for BAK patients, $p=0.8162$). Similarly, the mean preoperative ROM at L5–S1 for all patients (CHARITÉ and BAK) implanted

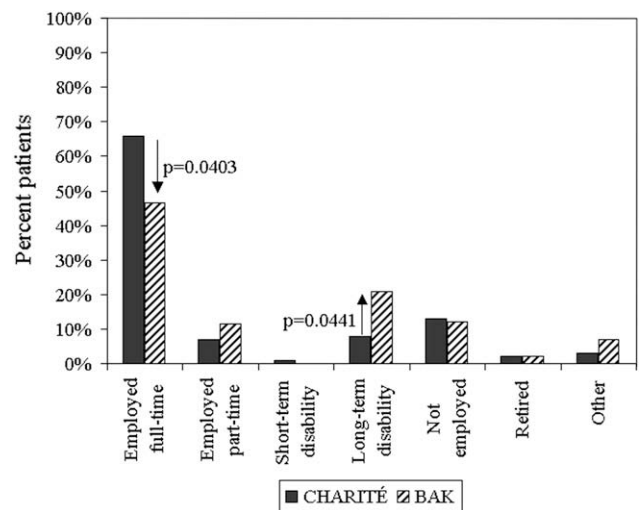


Fig. 5. Work and disability status at the five-year postoperative time point, for both treatment groups. In the CHARITÉ group, 65.6% patients were employed full-time vs. 46.5% in the BAK group. In addition, fewer CHARITÉ patients were on long-term disability (8.0%) compared with those in the BAK group (20.9%).

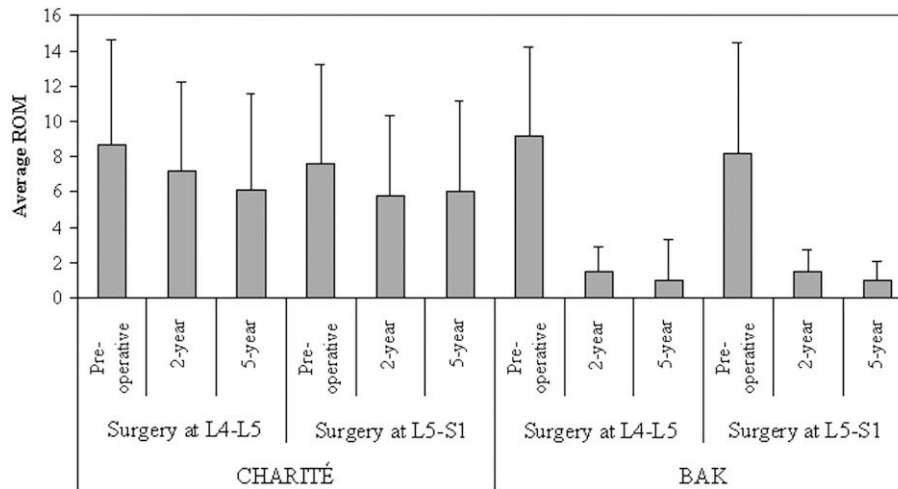


Fig. 6. Index-level ROM for CHARITÉ and BAK patients at the two- and five-year follow-up time points. No statistical difference in ROM was observed between any time point, for either CHARITÉ or BAK patients.

at the L5–S1 level was 7.8° and was not statistically different between the two groups (ROM=7.6° for CHARITÉ patients and 8.2° for BAK patients, $p=0.6793$).

For CHARITÉ patients implanted at L4–L5, the mean ROM at the index level did not statistically change from baseline to 2-years postoperative ($p=0.3351$). Similarly the mean change in ROM at index level from 2- to 5-years postoperative was not statistically significant ($p=0.1294$). However, the mean ROM change at index level from baseline to 5-years postoperative showed a trend toward statistical significance ($p=0.0539$). For CHARITÉ patients implanted at L5–S1, none of the mean changes in ROM were statistically significant (from baseline to two-year postsurgery: $p=0.3194$; from 2- to 5-year postsurgery: $p=0.2065$ and from baseline to five-year postsurgery: $p=0.2237$). Changes in mean index-level ROM for both L4–L5 and L5–S1 implanted BAK patients were statistically significant from baseline to either 2- or 5-year follow-ups (for L4–L5 implanted BAK patients, from baseline to 2-year postsurgery: $p=0.0076$; from baseline to 5-year postoperative: $p=0.0003$; for L5–S1 implanted BAK patients, from baseline to 2-year postsurgery and from baseline to 5-year postsurgery: $p\leq 0.0001$). However, there was no statistical difference between the mean index-level ROM changes from 2- to 5-year postsurgery for either L4–L5 or L5–S1 (L4–L5: $p=0.6413$; L5–S1: $p=0.2593$).

Adjacent-level range ROM was not statistically different from baseline to two- or five-year postoperative, for CHARITÉ or BAK patients, regardless of the implanted level (eg; L4–L5 or L5–S1).

Segmental translation and changes in disc height

No differences were observed between CHARITÉ and BAK patients in terms of changes in disc height. At the 5-year time point, disc height changes decreased by 0.7 mm for both CHARITÉ and BAK patients ($p=0.9827$).

Vertebral translation ranged from 0.4 mm to 0.8 mm in CHARITÉ-implanted L4–L5 and L5–S1 sites, respectively. This translation was within normal range and significantly lower than the 4.5 mm limit for clinical instability, as suggested by White and Panjabi [22]. The average vertebral translation was observed at 0.1 mm in BAK-implanted patients at both L4–L5 and L5–S1 sites.

Longitudinal ossification

Clinically relevant longitudinal ossification was defined as ossification in the disc space that would impact motion. The definition of motion was based on 2 different sources, the FDA guidelines (5°) or the ProDisc L SSE (3°).

Within the CHARITÉ group and using the 5° cut-off point to determine motion, 17 (18.9%) cases showed lack of motion and a rating ≥ 3 on the longitudinal classification system. Nine cases were implanted at L4–L5 (34.6% of all L4–L5 cases), and 8 were implanted at L5–S1 (12.5% of all L5–S1 cases). Using the 3° cut-off point to determine motion, 14 CHARITÉ cases (15.5%) met both criteria. Half of these were at the L4–L5 level (26.9% of all CHARITÉ patients operated at L4–L5) and the other half, at the L5–S1 level (10.9% of all CHARITÉ patients implanted at L5–S1).

Additional statistical analyses

The impact of missing data on study results was evaluated by comparing the clinical outcomes of the 5-year completers with that of all patients lost to follow-up (LTF). Baseline VAS and ODI comparisons showed no statistically significant difference in the BAK group between 5-year completers and LTF (ODI: $p=0.5229$; VAS: $p=0.3940$). In the CHARITÉ group, both ODI and VAS values were statistically higher in the LTF group than in the 5-year completers group (ODI: LTF=52.6 vs. 5-year completers=48.0, $p=0.0134$; VAS: LTF=73.7 vs. 5-year

completers = 69.7 $p=0.0496$). However, at the 2-year time point, there was no statistically significant difference between 5-year completers and LTF in either CHARITÉ or BAK groups, for VAS or ODI values.

The clinical outcomes at baseline and at the 2-year time points were also compared for all patients included in the 9 participating sites vs. those included in the 6 sites that declined participation. Baseline VAS and ODI comparisons showed no statistically significant difference in the BAK group between participating and nonparticipating sites (ODI: $p=0.7716$; VAS: $p=0.8316$). In the CHARITÉ group, both ODI and VAS values were statistically higher in the nonparticipating group than in the participating group (ODI: nonparticipating sites=55.5 vs. participating sites=49.2, $p=0.0049$; VAS: nonparticipating site=76.0 vs. participating site=70.9 $p=0.0400$). However, at the 2-year time point, there was again no statistically significant difference between patients from participating and nonparticipating sites in either CHARITÉ or BAK groups, for VAS or ODI.

The LOCF analysis was completed by using available results for subjects with 60-month follow-up and the last available result from the 2-year PMA subjects that did not have 60-month follow-up results available. A Blackwelder's test was used to evaluate the noninferiority of the CHARITÉ to the BAK assuming a Δ of 0.10. This test provided a $p=0.0009$.

Using Bayesian statistical methodology, the probability for the CHARITÉ group to achieve superiority ($\Delta=0$) vs. the BAK group was 78.5%, while the probabilities of the CHARITÉ group to be noninferior to the BAK group using a Δ of 5%, 10%, and 15% was 92.3%, 98.0%, and 99.7%, respectively.

Discussion

The purpose of this study was to provide 5-year clinical follow-up data on patients enrolled in a previously reported 2-year IDE randomized clinical study [10,11]. Previous publications already reported on the safety and efficacy of the CHARITÉ total disc replacement as compared with BAK with autograft, for the treatment of DDD at L4–L5 or L5–S1. This 5-year study confirms that the hypothesis validated at the 2-year time point was maintained throughout the 5-year follow-up period. An example of radiographic views of the implant at L5–S1 at the 5-year time point is shown in Fig. 7.

Clinical improvements were significant in both the CHARITÉ and BAK groups. The MCID for ODI, VAS, and SF-36 (PCS) were met early postoperatively and at the two- and five-year time points, improvements in ODI, VAS, and SF-36 (PCS) were twice as high as their respective MCID values.

A subset of the entire IDE patient population was available at 5 years postsurgery and thus has been included herein. A significant drop in the eligible patient population was initially due to site withdrawals from the study. All of the 14 sites involved in the initial 2-year CHARITÉ IDE trial were invited to participate in the 5-year follow-up; however, 6 of the 14 sites declined continuation. It is important to mention that all sites initially signed up for a 2-year study; they were therefore, under no obligation to pursue the investigation beyond the 2-year time point. There was no consistent reason why these sites declined participation; some sites did not have adequate clinical research support anymore, whereas others were involved in new clinical studies. The impact of this loss was analyzed



Fig. 7. Radiographic views of a subject implanted at L5–S1, at the 5-year follow-up time point, in (Left) lateral, (Middle) flexion, and (Right) extension.

statistically by comparing the clinical outcomes at baseline and 2-year postsurgery of participating vs. nonparticipating sites. This analysis showed that, although patients in the nonparticipating sites started with a statistically higher pain and disability rate than patients from the participating sites, at the 2-year time point, both groups were statistically similar. Additional loss to follow-up included 96 patients who did not respond to requests from their sites, which, via a third party investigator, attempted contact in multiple ways. The follow-up rate for this study thus reached 57% of the randomized eligible population (133/233) or 44% of the randomized IDE population (133/304).

The sample size in this analysis included 90 patients in the CHARITÉ group and 43 patients in the BAK group. These numbers are relatively small, a consequence of using a 2-year study design for a 5-year follow-up analysis. The initial sample size for the 2-year IDE trial was determined using the Blackwelder methodology [23] and was calculated on the basis of a 10% dropout rate. Although this rate is appropriate for a 2-year follow-up investigation, it underestimates the dropout expected from a 5-year study, as seen in other comparable long-term studies [24,25]. However, from a statistical standpoint, although small, the sample sizes used herein were sufficient to provide confidence of noninferiority between groups.

To determine whether this 133 patient population was indeed representative of the 304 initially enrolled, different statistical analyses were conducted. When analyzing the clinical outcomes of 5-year completers (for both CHARITÉ and BAK) vs. patients lost to follow-up (LTF), it was observed that at 2-year postsurgery, these 2 groups were statistically similar, despite the fact that CHARITÉ patients in the LTF group started at a disadvantage with statistically greater VAS and ODI values. In addition, the 5-year completers CHARITÉ vs. LTF CHARITÉ groups were evaluated, at 2-years, for full-time work and disability status. At that time point, no statistical difference was observed between these two patient populations (full-time employment: $p=0.2448$; disability: $p=0.7906$), further indicating that the 5-year completers CHARITÉ group was likely to represent the entire cohort, at the end of the 5-year time point.

The results of this study provided a high degree of confidence that the CHARITÉ group was non-inferior to the BAK group. To further strengthen this finding, a last observation carried forward (LOCF) analysis was completed and resulted in a $p=0.0009$, and a Bayesian statistical analyses provided a 99.5% probability for the CHARITÉ group to being non-inferior to BAK (at a $\Delta=0.1$). These additional statistical analyses further confirmed noninferiority of the CHARITÉ group as compared with the BAK group.

It is possible that among the lost to follow-up patients, there are those who have brought claims for personal injuries rather than continuing in the study. Due to a lack of information and other constraints, we are not presently able to determine the significance of these patient's outcomes to the study.

Blumenthal et al. previously discussed the rationale and scientific validity for selecting a BAK control group. In this present study, the BAK group had 98% radiographic fusion, as determined by longitudinal ossification measurements from independent radiologists. Clinical fusion success was at 88.4%, as defined by the surgeons during the patients' 5-year follow-up examinations. Radiographic and clinical fusion rates were different as determined by 2 independent medical professionals, thus guarding against potential bias. The high radiographic and clinical fusion rates observed herein for the BAK cohort alleviates concerns that this control may set the bar too low to truly evaluate a novel technology. Interestingly, fusion and reoperation rates from this BAK group were not very different from those published using the LT-Cage with rhBMP-2 in an ALIF procedure. Excluding reoperations, the rhBMP-2 study resulted in a 94.5% fusion rate (as determined with computed tomography scans) compared with the 88.4% clinical fusion rates (or 98% radiographic fusion rate) observed with BAK. In terms of reoperations, the rhBMP-2 group had a 10.4% failure rate (30 out of 288) at 2 years, compared with 16.3% failure rate (6 out of 43) at 5 years, for BAK [26,27]. Thus, the results from our BAK group were not very different from some of the best results published to date for anterior lumbar fusion.

When analyzing the 5-year CHARITÉ and BAK groups, establishing baseline similarities was essential to confirm adequate randomization. Demographic baseline data was therefore analyzed and found to be statistically equivalent, thus confirming an effective randomization process with no evidence of selection bias between CHARITÉ and BAK treatments. Although age, weight, activity level, and work status were similar across groups, there was a slight trend of increased incidence of prior lumbar surgery in the CHARITÉ group. Prior lumbar surgery was defined as minor interventions such as discectomies or laminectomies. Although there is very little evidence on the impact of prior laminectomies and/or discectomies on fusion outcomes, a recently published report presented high fusion rate in patients with prior laminectomies [28], thus indicating that these prior surgeries may not affect fusion outcomes. In any case, the patient population most affected by prior surgery was the CHARITÉ group and any clinical disadvantage resulting from prior surgery would have accrued to the CHARITÉ results.

The clinical impact of performing either a total disc replacement with the CHARITÉ artificial disc or an ALIF with a BAK cage and iliac crest autograft in this patient population was demonstrated by significant improvements in pain and disability scores. Only patients with debilitating back pain were included in this study. Median preoperative ODI values reached 48.0 pts in the CHARITÉ group and 52.0 pts in the BAK group (no statistical difference). Similarly, median VAS for both CHARITÉ and BAK groups were exactly 70.0. Comparatively, burst fractures and metastatic bone tumors in the spine were both previously given ODI values of approximately 60 [29]. An additional comparative figure was provided in the review paper by

Fairbank et al. [30] in which ODI scores were categorized by disease: chronic low back pain had the third worst ODI score with 43.3 pts and was only 4.74 pts lower than bone metastasis. These severely disabled patients had all undergone at least 6 months conservative care to meet the inclusion criteria of the study. On average, conservative care was provided to the CHARITÉ patients for 34.2 months (median: 24.0 months) and to the BAK patients for 26.3 months (median: 21.0 months). Opportunities for nonoperative care for these patients had therefore been exhausted.

In both the CHARITÉ and the BAK groups, ODI and VAS improvements exceeded MCID thresholds by the three-month time point and reached nearly twice the MCID threshold by two and five years. These results showed that, in both groups, pain and functional disability was improved for most of the patients at early time points and remained favorable throughout the 5-year follow-up.

Additional surgery, whether for index-level or at adjacent-level degenerative disc disease, were relatively low for both CHARITÉ and BAK groups compared with prior reports with BAK [31,32]. Degenerative disc disease, defined as requiring secondary surgery, affected 1.1% CHARITÉ patients and 4.7% BAK patients, these numbers being too low to draw statistical conclusions.

Work statuses at two- and five-year follow-ups were also evaluated. A statistically greater number of patients were employed full-time in the CHARITÉ group as compared with the BAK group. In addition, a statistically smaller number of patients in the CHARITÉ group were on long-term disability, as compared with BAK patients.

ROM at index and adjacent levels did not change statistically from the 2- to the 5-year time points, for either CHARITÉ or BAK patients. Additional analyses were conducted to evaluate any relationship between index-level ROM and clinical outcome (VAS and ODI), but in both cases, no correlation was found. In addition, at the 5-year time point, segmental translation and disc height were equivalent between both groups. Segmental translation was ranging from 0.4 and 0.8 mm in CHARITÉ-implanted patients. To place this result into context, the following published data on segmental translation need to be considered. Percy and Posner et al., in two separate publications, described a normal range of segmental translation from 2.0 to 2.8 mm [33,34]. Panjabi further defined clinical instability as a segmental translation greater than 4.5 mm [35]. These values are significantly greater than the segmental translation observed in CHARITÉ cases, thus suggesting normal segmental translation for those patients.

Conclusions

Results of this five-year, prospective, randomized multicenter study are consistent with the two-year reports of non-inferiority of CHARITÉ artificial disc vs. ALIF with BAK and iliac crest autograft. No statistical differences were found in safety and clinical success outcomes between groups. This

study represents the largest and longest randomized, prospective, multicenter arthroplasty trial performed to date.

Acknowledgment

We thank Dr. Chantal E. Holy and Dan Stoutenburgh, employees of DePuy Spine, for editorial support.

References

- [1] Link HD. Letter to the Editor concerning "Charite total disc replacement: clinical and radiographical results after an average follow-up of 17 years" (M. Putzier et al.). [comment]. *Eur Spine J* 2006;15:514–7.
- [2] Buttner-Jan K. Letter to the Editor concerning "Charite total disc replacement: clinical and radiographical results after an average follow-up of 17 years" (M. Putzier et al.). *Eur Spine J* 2006;15:510–3.
- [3] Cunningham BW, Gordon JD, Dmitriev AE, Hu N, McAfee PC. Biomechanical evaluation of total disc replacement arthroplasty: an in vitro human cadaveric model. *Spine* 2003;28:S110–7.
- [4] Cunningham BW, Dmitriev AE, Hu N, McAfee PC. General principles of total disc replacement arthroplasty: seventeen cases in a non-human primate model. *Spine* 2003;28:S118–24.
- [5] Griffith SL, Shelokov AP, Buttner-Jan K, Lemaire JP, Zeegers WS. A multicenter retrospective study of the clinical results of the LINK SB Charite intervertebral prosthesis. The initial European experience. *Spine* 1994;19:1842–9.
- [6] Cinotti G, David T, Postacchini F. Results of disc prosthesis after a minimum follow-up period of 2 years. *Spine* 1996;21:995–1000.
- [7] Lemaire JP, Carrier H, SariAli e, Skalli W, Lavaste F. Clinical and radiological outcomes with the Charite artificial disc: a 10-year minimum follow-up. [erratum appears in *J Spinal Disord Tech*. 2006 Feb;19(1):76 Note: Sari Ali, El-Hadi (corrected to SariAli, El-hadi)]. *J Spinal Disord Tech* 2005;18:353–9.
- [8] Putzier M, Funk JF, Schneider SV, et al. Charite total disc replacement—clinical and radiographical results after an average follow-up of 17 years. [see comment]. *Eur Spine J* 2006;15:183–95.
- [9] David T. Long-term results of one-level lumbar arthroplasty: minimum 10-year follow-up of the CHARITE artificial disc in 106 patients. *Spine* 2007;32:661–6.
- [10] Blumenthal S, McAfee PC, Guyer RD, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. [erratum appears in *Spine*. 2005 Oct 15;30(20):2356]. *Spine* 2005;30:1565–75.
- [11] McAfee PC, Cunningham B, Holsapple G, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. *Spine* 2005;30:1576–83.
- [12] Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160–4.
- [13] Geisler FH. Surgical technique of lumbar artificial disc replacement with the Charite artificial disc. *Neurosurgery* 2005;56(Suppl. 57).
- [14] Kuslich SD, Ulstrom CL, Griffith SL, Ahern JW, Dowdle JD. The Bagby and Kuslich method of lumbar interbody fusion. History, techniques, and 2-year follow-up results of a United States prospective, multicenter trial. *Spine* 1998;23:1267–78.
- [15] Zhao K, Yang C, Zhao C, An KN. Assessment of non-invasive intervertebral motion measurements in the lumbar spine. *J Biomech* 2005;38:1943–6.
- [16] Nathan H. Osteophytes of the vertebral column: an anatomical study of their development according to age, race, and sex with

- considerations as to their etiology and significance. *J Bone Joint Surg Am* 1962;44:243–68.
- [17] McAfee PC, Fedder IL, Saiedy S, Shucosky EM, Cunningham BW. SB Charite disc replacement: report of 60 prospective randomized cases in a US center. *J Spinal Disord Tech* 2003;16:424–33.
- [18] Food and Drug Administration. Guidance document for the preparation of IDEs for spinal systems. 2000.
- [19] Food and Drug Administration. PRODISC[®]-L total disc replacement—P050010: summary of safety and effectiveness data; 2006.
- [20] Hagg O, Fritzell P, Nordwall A. Swedish Lumbar Spine Study Group. The clinical importance of changes in outcome scores after treatment for chronic low back pain. [see comment]. *Eur Spine J* 2003;12:12–20.
- [21] Polly DW Jr, Glassman SD, Schwender JD, et al. SF-36 PCS benefit-cost ratio of lumbar fusion comparison to other surgical interventions: a thought experiment. *Spine* 2007;32(Suppl 6).
- [22] White AW, Panjabi MM. *Clinical biomechanics of the spine*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 1990.
- [23] Blackwelder WC, Chang MA. Sample size graphs for “proving the null hypothesis. *Control Clin Trials* 1984;5:97–105.
- [24] Glaser J, Stanley M, Sayre H, Woody J, Found E, Spratt K. A 10-year follow-up evaluation of lumbar spine fusion with pedicle screw fixation. *Spine* 2003;28:1390–5.
- [25] Kuslich SD, Danielson G, Dowdle JD, et al. Four-year follow-up results of lumbar spine arthrodesis using the Bagby and Kuslich lumbar fusion cage. *Spine* 2000;25:2656–62.
- [26] Burkus JK, Gornet MF, Dickman CA, Zdeblick TA. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. *J Spinal Disord Tech* 2002;15:337–49.
- [27] Food and Drug Administration. INFUSE[™] bone graft / LT-CAGE[™] lumbar tapered fusion device. Summary of effectiveness and data P000058; 2002.
- [28] Badawy WS, El Masry MA, Radwan YA, El Haddidi TT. Results of instrumented posterolateral fusion in failed back surgery. *Int Orthop* 2006;30:305–8.
- [29] Wood K, Buttermann G, Mehdod A, et al. Operative compared with nonoperative treatment of a thoracolumbar burst fracture without neurological deficit. A prospective, randomized study. [see comment; erratum appears in *J Bone Joint Surg Am*. 2004 Jun;86-A(6):1283]. *J Bone Joint Surg Am* 2003;85-A:773–81.
- [30] Fairbank JC, Pynsent PB. The Oswestry disability index. *Spine* 2000;25:2940–52.
- [31] Button G, Gupta M, Barrett C, Cammack P, Benson D. Three- to six-year follow-up of stand-alone BAK cages implanted by a single surgeon. *Spine J* 2005;5:155–60.
- [32] Pavlov PW, Spruit M, Havinga M, Anderson PG, van LJ, Jacobs WC. Anterior lumbar interbody fusion with threaded fusion cages and autologous bone grafts. *Eur Spine J* 2000;9:224–9.
- [33] Percy MJ. Stereo radiography of lumbar spine motion. *Acta Orthop Scand* 1985;212(Suppl):1–45.
- [34] Posner I, White AA III, Edwards WT, Hayes WC. A biomechanical analysis of the clinical stability of the lumbar and lumbosacral spine. *Spine* 1982;7:374–89.
- [35] Panjabi M, Malcolmson G, Teng E, Tominaga Y, Henderson G, Serhan H. Hybrid testing of lumbar CHARITE discs versus fusions. *Spine* 2007;32:959–66.