

Does Vertebral Endplate Morphology Influence Outcomes in Lumbar Total Disc Arthroplasty?

Part II: Clinical and Radiographic Results as Evaluated Utilizing the Vertebral Endplate Yue-Bertagnoli (VEYBR) Classification

Matthew E. Oetgen, MD,^a James J. Yue, MD,^a Jorge J. Jaramillo-de la Torre, MD,^a
and Rudolf Bertagnoli, MD^b

ABSTRACT

Study Design

This study was a prospective study with a minimum patient follow-up of 2 years.

Objective

The purpose of this study was to evaluate the clinical usefulness of a vertebral endplate classification system (VEYBR) in predicting outcomes following lumbar arthroplasty.

Background

In the present study, our previously described endplate classification system was evaluated to determine its clinical usefulness in patients undergoing lumbar arthroplasty.

Methods

The patient cohort in this study consisted of 80 patients who had been enrolled in the US FDA ProDisc clinical trial. Radiographs were classified using the VEYBR classification. The preoperative categories (Types I to V) were then correlated with the patients' visual analogue scores (VAS) and Oswestry Disability Index (ODI) scores and radiographic outcomes at an average follow-up point of 28 months.

Results

The rank order of total change in VAS based on preoperative VEYBR classification was Type IV, III, I II, and V, with Type IV having the greatest improvement in VAS and Type V having the least improvement. The rank order of total change in ODI was Type IV, II, III, I, and V. We found no differences in clinical outcomes among the 5 vertebral endplate types. Type II endplates had least optimal sagittal positioning.

Conclusions

Although not statistically significant, there was a strong trend for Type V endplates to have the least improvements in VAS and Oswestry clinical outcome scores. Knowledge and use of the endplate classification system did lead to consistent implant placement across endplate classes which may indicate the usefulness of this classification system in preoperative planning, especially for physicians in the "learning curve" phase of this procedure.

Level of Evidence

Case series (Level IV).

Key Words: Lumbar, endplate morphology, total disc replacement, clinical and radiographic outcomes. *SAS Journal*. Spring 2008;2:101-106. DOI: SASJ-2007-0119-RR

^aDepartment of Orthopaedics and Rehabilitation, Yale University School of Medicine, New Haven, Connecticut; ^bSt. Elizabeth's Klinikum, Straubing, Germany

Address correspondence to James J. Yue, MD, Department of Orthopaedics and Rehabilitation, Yale University School of Medicine, P.O. Box 208071, New Haven, CT 06520-8071 (e-mail: james.yue@yale.edu)

None of the authors has received direct research support for this project. James Y. Yue, MD, and Rudolf Bertagnoli, MD, are consultants to Synthes Spine.

Institutional Review Board approval was obtained from the Yale Human Investigation Committee.

INTRODUCTION

The goals of any classification system are to allow patients to be reliably compared within and across physicians, assist in preoperative planning for treatment purposes, and predict patient outcome. Until now, there has been no published classification system for patients considered for total disc arthroplasty (TDA). In Part I of this study we introduced a classification for patients undergoing lumbar TDA based on the endplate morphology assessed on preoperative lateral radiographs.¹ Our preliminary research demonstrated the classification system was reliable and valid. In our first study we did not assess clinical reliability. The goal of the present research was to assess clinical reliability with regards to this novel classification system.

Outcomes for patients undergoing TDA have been evaluated in numerous studies.²⁻⁸ Factors affecting outcome have been ascertained by retrospectively analyzing large clinical trials. Two important factors reported in the outcomes of patients following TDA are the implant position and postoperative range of motion.^{4,6} Traditional factors that have been shown to negatively affect outcome in lumbar spine surgery, such as smoking and age, have been shown to be less important in the outcome of patients undergoing TDA.^{9,10} It would be very useful to have a way to predict outcome based on preoperative information instead of postoperative evaluation, such as postoperative position of device. This information may lead to better decision making in terms of patient and implant selection for this relatively new procedure.

In order to assess the clinical reliability and usefulness of our new classification system for TDA, we analyzed patient outcomes based on the preoperative classification of the patients' vertebral endplates. We attempted to determine whether our new classification system was useful in predicting clinical outcome in patients undergoing TDA.

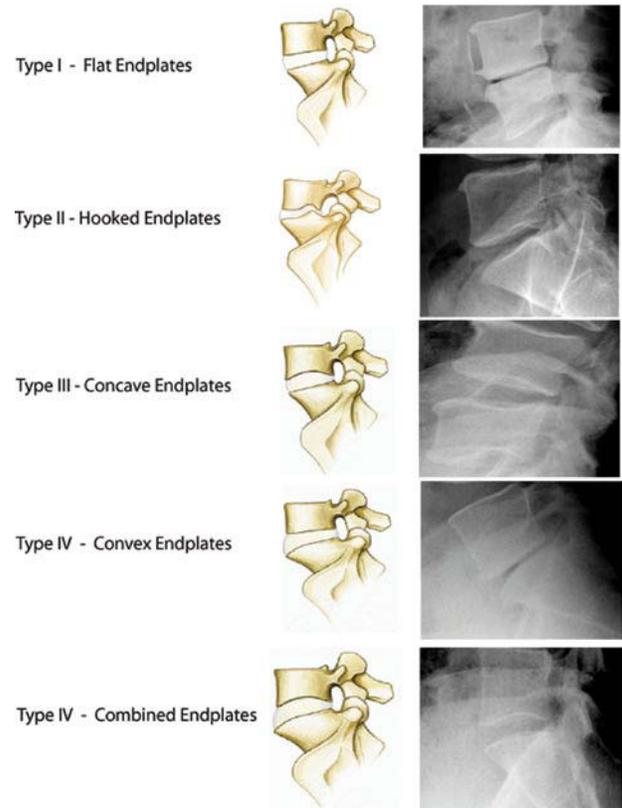
METHODS

As previously reported, the study population consisted of 80 consecutive patients (119 disc levels) undergoing total disc arthroplasty as participants in a prospective clinical and radiographic outcome analysis of the ProDisc-L (Synthes, West Chester, Pennsylvania) total disc prosthesis. All patients underwent single- or bi-segmental total disc replacement utilizing the named prosthesis by a single orthopaedic spine surgeon.

Endplate Classification

As part of the original study evaluating the reliability of our new endplate classification, each surgical level was evaluated multiple times by 3 different physicians and classified as Types 1–5 (Figure 1). A consensus classification of each level was then assigned after each surgical level was evaluated and agreed upon by the 3 evaluators as a group. In patients who underwent multiple-level disc replacements, their overall classification was recorded as a single value for data analysis.

Figure 1.



Five types of lumbar endplates: Type I - Flat endplate; Type II - Posterior hooked endplate; Type III - Concave endplate; Type IV - Convex endplate; Type V - Combined endplates.

If the patient had 2 levels that were classified as Type 1, their overall classification was recorded as Type 1. If the patient had 1 endplate that was classified as Type 1 and another endplate that was something other than Type 1, the patient's overall classification was recorded as the higher value. If the patient had multiple levels evaluated as other than Type 1, they were given an overall classification of Type 5.

Patient Outcomes

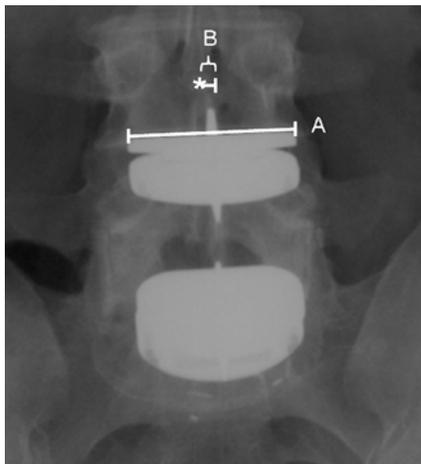
Patients were followed postoperatively at regular intervals (6, 12, 18, and 24 weeks) and evaluated both preoperatively and postoperatively by independent observers using multiple outcome assessment tools including the Oswestry Disability Index (ODI) score and a visual analogue pain score (VAS).¹¹

Disc Position

The postoperative radiographs of each patient were evaluated for the position of the artificial disc. Each operative level was analyzed according to the method described by McAfee et al., in which the distances from perfect placement on both the postoperative anteroposterior (AP) and lateral radiographs are measured.⁶ Measurements were made by hand. Midline marking in the AP plane was determined to be a point equidistant from the medial border of the pedicles. The midline in the lateral plane was measured from the most

anterior and posterior vertebral cortical edges. The center keel was used to determine the center of the implant. Magnification markers were placed at time of radiographic imaging. Perfect placement of the artificial disc was defined as at the midline of the vertebral body in the coronal plane and 2 mm posterior to the midline of the vertebral body in the sagittal plane. The disc position was classified as ideal, suboptimal, or poor based on the measured position of the artificial disc. Ideal placement was defined as within 3 mm of perfect placement on both the AP and lateral radiograph. Suboptimal placement was defined as 3–5 mm from perfect placement on either the AP or lateral radiograph. Poor placement was defined as > 5 mm from perfect placement on either the AP or lateral radiograph. The measured value of deviation from perfect placement of each level was corrected for radiographic magnification error (Figures 2A and 2B).

Figure 2A.



AP radiograph of bi-segmental TDA demonstrating the measurements used to determine the placement of the artificial disc in relation to the true midline of the vertebral body. Line A represents the measured length of the artificial disc baseplate. Line B represents the distance between the midline of the artificial disc and the midline of the vertebral body (marked with *). The midline of the vertebral body was determined by measuring the width of the vertebral body at the endplate and using the midpoint of this measurement. To control for differences in radiographic magnification, the measurement of distance from the center was controlled using the known length of the artificial baseplate with the formula: **True Distance from Center = B - (B * (A / actual implant size) - 1)**.

Statistical Analysis

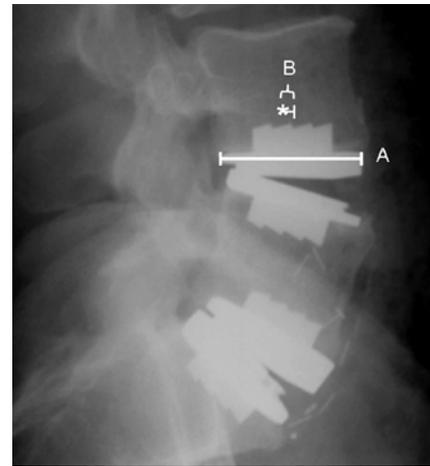
Statistical analysis was performed using Minitab Statistical Software version 13.1 (State College, Pennsylvania). Continuous data were compared across endplate classifications using one-way ANOVA tests. Categorical data were compared using a chi-square test. Statistical significance was defined as $P < .05$.

Data Analysis

The absolute deviations of the implanted device from perfect position on both the postoperative AP and lateral radiographs were compared using a one-way ANOVA according to the

preoperative endplate classifications for each operative vertebral level.

Figure 2B.



Lateral radiograph of the same bi-segmental TDA showing the radiographic measurements utilized.

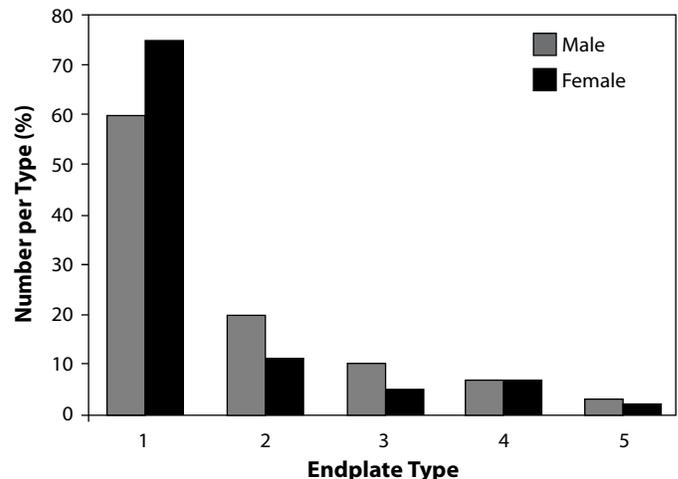
Next, each patient’s clinical outcomes (Oswestry and VAS scores) were compared according to the position of the implanted disc (ideal, suboptimal, or poor) using a one-way ANOVA test.

Finally, a chi-square analysis was performed to determine the effects of preoperative endplate classification on the position (ideal, suboptimal, or poor) of the implanted device.

RESULTS

Eighty patients underwent TDA during the study period, 41 patients had single-level procedures and 39 patients had multilevel procedures. The mean follow-up for the group was 28 months (range 10–47 months). Gender and general medical and smoking status did not correlate with endplate

Figure 3.



Gender and endplate type.*

*No statistical difference between males and females for each endplate type

Table 1. Endplate Type, Smoking, and Medical Status

Endplate Type (Number of Subjects)	Age (Avg. of group)	# Smokers (%)	Depression/ Fibromyalgia/ Bipolar Dz (No. per group)
1(43)	38	6(33)	11
2(17)	39	0 (0)	4
3(9)	40	3(33)	0
4(7)	37	3(75)	0
5(4)	40	3(75)	1

*Age: $P = .87$ for age between groups (ANOVA); Smk: $P = .004$; DFB: $P = .28$

Table 2. Frequency of Endplate Type

Endplate Type	Number of Levels (# of Total Endplates Evaluated)
Type I	78 (66%)
Type II	20 (17%)
Type III	11 (9%)
Type IV	7 (6%)
Type V	3 (2%)

type (Figure 3 and Table 1). The frequency of endplate type is listed in Table 2. Type 1 endplates were the most common and Type 5 the least common.

General Radiographic Data and Patient Outcome

Table 3 delineates the levels of surgery. L5-S1 was the most common level in single level cases. L4-5 was the most common level in 2-level cases. Average range of motion was 7.1 degrees. Average disc height was 10mm for operative levels. Clinical outcomes as measured by both VAS and Oswestry scores did not directly correlate with range of motion ($P > .2$ and $P > .7$, respectively). We experienced 3 cases of subsidence (3.75%). Clinical outcome was not significantly affected by cases of subsidence.

Table 3. Distribution of Operated Levels

Patient Group	L3-4	L4-5	L5-S1
Total Patients (n=80)	6	51	62
Single Level (n=41)	2	12	27
Multilevel (n=39)	4	39	35

Patient Outcome and Endplate Type

Overall, patients showed a statistically significant improvement in both their Oswestry disability score and VAS score postoperatively (Table 4). We then analyzed patients' clinical outcomes in terms of their preoperative endplate classification, as shown in Table 5. Patients in each endplate category showed statistically significant outcome improvements as measured by both the ODI score and VAS score. We did not find any differences in preoperative, postoperative, or total change of these outcomes between the endplate classes. The rank order of total change in VAS was Type IV, III, I II, and V, with Type IV having the greatest improvement in VAS and Type V having the least improvement. The rank order of total change in ODI was

Type IV, II, III, I, and V.

Implant Position and Endplate Type

The absolute deviation in the position of each artificial disc in both the coronal plane and the sagittal plane from the perfect placement is shown in Table 6. The endplate type did not appear to affect the position of the implanted disc, as we found no difference in implant position when compared across endplate types. Type II endplates had the least favorable sagittal alignment of the 5 endplate types.

Implant Position and Patient Outcome

Patient outcomes, as indicated by the VAS pain score and Oswestry disability score, as factors of implant position (ideal, suboptimal, or poor) are shown in Table 7. We found that patients in whom implants were placed in a suboptimal position had better outcomes when compared to patients identified as having ideally positioned implants. We found no statistically significant difference between patients' outcomes with poorly positioned implants as compared to ideally or suboptimally positioned implants.

Endplate Type and Implant Position

The association between endplate type and implant position was evaluated by comparing the postoperative implant position (ideal versus suboptimal and poor) in Type I endplates and the implant position in Types II–V endplates (Table 8). We found no statistical difference in the number of suboptimal and poor implant positions postoperatively in the Type I compared to Types II–V endplates.

DISCUSSION

We have previously shown our new classification system for preoperative assessment of vertebral endplates prior to TDA to be reliable and valid. The results of the second part of the evaluation of this classification system, the clinical evaluation, suggest preoperative recognition of differing endplate and appropriate intraoperative re-shaping of dysmorphic endplates may equalize clinical outcomes among the 5 different endplate morphologies. While we found the overall patient outcomes significantly improved from preoperative levels after TDA, there were no differences in postoperative outcomes when the patients were stratified by preoperative endplate type (Tables 4 and 5).

Our new classification system was developed to account for

Table 4. Patient Outcome Data

	Preop	Postop	P value
Oswestry Disability Score	67%	29%	<0.01
VAS Pain Score	77	32	<0.01

variations in the morphology of the vertebral endplates which may affect the placement of artificial disc base plates if not addressed at the time of surgery. We hypothesized that if the

Table 5. Patient Outcome an Endplate Classification

Patient Outcome	Type 1	Type 2	Type 3	Type 4	Type 5	P value
Preop VAS	79.28	70.62	79.08	76.46	70.47	.36
Postop VAS	35.12	27.77	32.56	21.42	32.77	.78
Total Change VAS	-44.16	-42.85	-46.53	-55.04	-37.70	.90
Preop Oswestry	68.76	64.00	68.20	64.67	66.00	.72
Postop Oswestry	32.59	23.56	30.40	16.33	34.50	.44
Total Change Oswestry	-36.17	-40.44	-37.80	-48.33	-31.50	.76

artificial disc placement was affected by these morphologic abnormalities, patient outcome may be affected by suboptimal placement of the implant. When we looked at the average deviation (in the coronal and sagittal planes) from midline of the placement of our implants, we found no difference when stratified by preoperative endplate type (Table 6). In addition, there was no difference in the number of poorly positioned implants in Type I endplates when compared to Types II–V endplates, further validating the limited interaction found between endplate type and implant position (Table 8). This uniform placement of our artificial discs, despite preoperative endplate type, may be the reason we found no clinical outcome difference between types.

Table 6. Implant Position and Endplate Classification

Endplate Class	Coronal Deviation (mm)	Sagittal Deviation (mm)
Type I	1.18	2.69
Type II	0.89	2.98
Type III	1.59	2.85
Type IV	0.98	2.45
Type V	1.45	2.42

*Age: $P = .87$ for age between groups (ANOVA); Smk: $P = .004$; DFB: $P = .28$

To date there are no definitively identified predictors of good clinical outcomes following TDA. Bertagnoli and Kumar correlated their patient outcomes following lumbar disc replacement with a number of preoperative patient factors, including number of affected disc spaces, degree of facet and disc degeneration, adjacent level disease, and instability.² Although they reported a stratification system of patients based on these factors which correlated with patient outcome from TDA, their outcome measures of “excellent, good, fair, and poor” were not well defined, limiting the clinical application of this system. Huang and colleagues demonstrated a statistically significant correlation between postoperative radiographically measured range of motion (ROM) and clinical outcome.⁴ They found higher ROMs were associated with improved clinical

Table 7. Implant Position and Patient Outcome

Implant Position	VAS	ODI
Ideal	39.57 *	35.1% *
Suboptimal	22.91 *	21.1% *
Poor	34.39	34.0%
P-value	.036	.032

*Indicates statistically different means based on Tukey’s post-hoc analysis

outcomes; however, they made no attempt to determine factors which led to improvements in ROM, again limiting the clinical usefulness of this information.

Table 8. Endplate Class and Implant Position

Endplate Class	Ideal Implant Position	Suboptimal and Poor Implant Position
Type I	48	30
Types II - V	23	17

Note: Chi-squared analysis for difference in implant position, $P = .671$.

Regan and colleagues studied the effect of surgical volume on patient outcome after TDA.⁷ Although the more experienced physicians had lower operative times and length of stay, there was no difference in patient outcome or adverse events. They defined “experienced” surgeons versus non-experienced surgeons as those who had performed at least 15 procedures. This may have been a relatively small number of cases to differentiate experienced versus not experienced, and the results may reflect the surgeons’ increased comfort with the procedure rather than a better understanding of the procedure. With an even greater number of cases performed improved understanding of the procedure would be expected and an improvement in patient outcome would likely ensue.

Finally, McAfee et al. correlated clinical outcome with postoperative positioning of the implant. They found implants positioned >5 mm from the midline in the coronal plane and > 5 mm from the perfect position in the sagittal plane (defined as 2 mm posterior to the vertebral body midline) were more likely to have a poor outcome.⁶ Implants positioned less than 5 mm from ideal placement did not differ in outcome.

Our data showed patients identified as having suboptimal (3–5 mm from perfect placement) implant positioning to have the best clinical outcome. This is at odds with the data found by McAfee et al.⁶ The major difference between these studies is the type of implant used. The ProDisc, which was evaluated in our study, is designed with a semi-constrained polyethylene core whereas the Charité (DePuy Spine, Raynham, Massachusetts) artificial disc (evaluated in the McAfee study) has a non-constrained polyethylene core design.¹² The difference in the design of the implants may necessitate different implant positioning on the vertebral endplates for optimal ROM and patient outcome, but this concept has not yet been evaluated.

Although we did not find our preoperative endplate classification to have a predictive value for patient outcome, a major bias inherent in our study may have misled our results. All of the procedures in this study were performed by a single surgeon who had a large clinical experience with total disc arthroplasty prior to performing the procedures evaluated in this study. In addition, the operative surgeon was the developer of the classification system which is being evaluated. This surgeon bias may have led to better outcomes than would have been expected with surgeons without experience with the classification system. We found no difference in average deviation from the midline of the implants across endplate types, contrary to what we hypothesized. One may interpret these results as actually confirming the clinical importance of our classification system, as a thorough knowledge of preoperative endplate type and attention to this factor during the surgical procedure could lead to uniform radiographic and clinical results across endplate type. Evaluating this classification system used by a group of surgeons in the “learning curve” phase of total disc implantation may more clearly show the predictive value of this classification system.

There were a number of weaknesses in this study. One such weakness was the inclusion of single- and 2-level disc replacement patients in the analysis. Although we attempted to limit the impact of this variable by using the worse of the classification when 2 levels were involved, it is unclear from the literature how multi-level disc replacements compare to single level replacements. Inclusion of only single-level replacements would have alleviated this weakness, at the cost of half of the patient population. The data reported were evaluated in the subgroup of single level patients, with no differences found from the data presented. Lastly, the authors do not believe it is possible to evaluate the effects of a properly placed implant at 1 level and an improperly placed implant at an additional level in a patient with multilevel ADR.

In summary, the results of the second part of the evaluation of this classification system, the clinical evaluation, suggest that preoperative recognition of differing endplate and appropriate intraoperative re-shaping of dysmorphic endplates may equalize clinical outcomes among the 5 different endplate morphologies. Our preoperative endplate classification may be very important in terms of patient outcome for surgeons in the “learning curve” phase of TDA and may be a vital tool in preoperative planning for this procedure. Additional studies exploring other ADR implants (eg, Charité, Flexicore, and others) and the effect of endplate morphology are warranted.

This manuscript was submitted October 14, 2007, and accepted for publication March 20, 2008.

REFERENCES

1. Yue J, Jaramillo-de la Torre J, Bertagnoli R. Does vertebral endplate morphology influence outcomes in lumbar disc arthroplasty? Part I: an initial assessment of a novel classification system of lumbar endplate morphology. *SAS Journal*. 2008;2:16-22.
2. Bertagnoli R, Kumar S. Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. *Eur Spine J*. 2002;11 Suppl 2:S131-136.
3. Blumenthal S, McAfee PC, Guyer RD, 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 I: evaluation of clinical outcomes.[erratum appears in *Spine*. 2005;30(20):2356]. *Spine*. 2005;30(14):1565-1575; discussion E1387-1591.
4. Huang RC, Girardi FP, Cammisia FP, Jr., Lim MR, Tropiano P, Marnay T. Correlation between range of motion and outcome after lumbar total disc replacement: 8.6-year follow-up. *Spine*. 2005;30(12):1407-1411.
5. Lemaire J-P, Carrier H, Sariali E-h, 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;19(1):76 Note: Sari Ali, El-Hadi [corrected to Sariali, El-hadi]]. *J Spinal Disord Tech*. 2005;18(4):353-359.
6. 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(14):1576-1583; discussion E1388-1590.
7. Regan JJ, McAfee PC, Blumenthal SL, et al. Evaluation of surgical volume and the early experience with lumbar total disc replacement as part of the investigational device exemption study of the Charite Artificial Disc. *Spine*. 2006;31(19):2270-2276.
8. Zigler JE. Clinical results with ProDisc: European experience and U.S. investigation device exemption study. *Spine*. 2003;28(20):S163-166.
9. Bertagnoli R, Yue JJ, Kershaw T, et al. Lumbar total disc arthroplasty utilizing the ProDisc prosthesis in smokers versus nonsmokers: a prospective study with 2-year minimum follow-up. *Spine*. 2006;31(9):992-997.
10. Bertagnoli R, Yue JJ, Nanieva R, et al. Lumbar total disc arthroplasty in patients older than 60 years of age: a prospective study of the ProDisc prosthesis with 2-year minimum follow-up period. *J Neurosurg Spine*. 2006;4(2):85-90.
11. Fairbank JC PP. The Oswestry Disability Index. *Spine*. 2000;25:2940-2952.
12. Petersilge CA. Lumbar disc replacement. *Semin Musculoskelet Radiol*. 2006;10(1):22-29.