SPINE Volume 30, Number 6, pp 675–681

Does the Number of Levels Affect Lumbar Fusion Outcome?

John J. Lettice, MD,* Thomas A. Kula, MD,* Richard Derby, MD,†‡
Byung-Jo Kim, MD, PhD,†§ Sang-Heon Lee, MD, PhD,† and Kwan Sik Seo, MD†

Study Design. Retrospective outcome measurement after circumferential reconstructive surgery with lumbar fusion in patients with chronic discogenic low back pain.

Objective. To examine the effect of the number of fusion levels on surgical outcomes in patients with chronic discogenic low back pain using provocative pressure-controlled diskography as a primary diagnostic tool.

Summary of Background Data. Although there is general agreement that construct length adversely affects arthrodesis success rates, the effect of the number of levels on lumbar fusion surgery outcome has not been reported. Previous fusion outcome studies have generally relied on magnetic resonance imaging or conventional diskography for diagnosis.

Methods. From 1994 through 2000, prospectively collected medical records of patients who underwent reconstructive lumbar spine surgery with confirmation of the pain generator by pressure-controlled diskography were retrospectively analyzed. Data were subdivided into 2 groups of patients. The first group, designated the short segment group, contained patients who underwent fusion at 1 or 2 levels. The second group, designated the long segment group, contained patients who underwent fusion at 3–5 levels. Surgical methods included circumferential reconstruction of the lumbar spine by either posterior or combined anterior and posterior approach. Surgeries included posterior decompression necessary to relieve documented regions of neural compression, combined with interbody arthrodesis at selective levels, augmented by posterior segmental spinal instrumentation and posterolateral arthrodesis. All patients completed a preoperative aquatic-conditioning program. Whenever possible, coexisting medical conditions were corrected or stabilized before surgery. A preoperative Short Form RAND 36-Item Health Survey (SF-36) was completed, and repeated at 1 and 2 years after surgery. The short and long segment groups contained 142 and 82 patients, respectively, who completed the preoperative SF-36 questionnaire completely.

Results. One hundred patients in the short segment group (*vide infra*) were available for 1-year follow-up, and 68 were available for 2-year follow-up. In the long segment group, 81 patients were available for 1-year follow-up, and 49 were available for 2-year follow-up. Mean ages were 41.0 and 47.6 years in the short and the long segment groups, respectively. The 2 groups did not differ significantly in gender, smoking habits, workers' compensation, or litigation ($P > 0.05$). In the short segment group, postoperative 1-year mean Physical Component Summary (PCS) and Mental Component Summary scores significantly improved ($P < 0.001$ and $P = 0.002$, respectively). Domains other than general health perceptions showed significantly improved 1-year follow-up scores ($P \leq 0.001$). Two-year follow-up scores showed significant improvement ($P < 0.001$ for physical function [PF], role function as limited by physical problems [RP], bodily pain [BP], social function [SF], and PCS). The vitality (VT) and role function as limited by emotional problems (RE) also improved ($P = 0.005$ and $P < 0.05$, respectively). In the long segment group, postoperative 1-year mean PCS scores improved significantly ($P < 0.001$), with some improvement in Mental Component Summary score ($P < 0.05$). The long segment group also showed significantly improved PF, RP, BP, and SF scores ($P < 0.001$). The VT and RE scores gave $P = 0.002$ and $P < 0.05$, respectively. Comparing preoperative and 2-year follow-up scores, PCS, PF, RP, BP, and SF showed significant improvement ($P < 0.001$), and the VT score gave $P < 0.01$. Mean difference in postoperative and preoperative scores for both groups did not show significant differences ($P > 0.05$), although the PF score showed differences in 1 and 2-year follow-up scores ($P = 0.048$ and $P = 0.068$, respectively).

Conclusions. When using strict patient selection criteria that include independent determination of pain generators *via* pressure-controlled diskography and completion of a preoperative conditioning program for improving general health status, the number of levels in reconstructive lumbar surgery may not significantly impact overall clinical outcome.

Key words: reconstructive surgery, lumbar fusion, chronic discogenic low back pain, number of levels, outcome, SF-36. **Spine 2005;30:675–681**

From the *Center for Spinal Deformity and Injury, Los Gatos, CA; †Spinal Diagnostics and Treatment Center, Daly City, CA; ‡Division of Physical Medicine and Rehabilitation, School of Medicine, Stanford University, Stanford, CA; and §Korea University Medical Center, Seoul, Korea.

Acknowledgment date: November 14, 2003. First revision date: January 27, 2004. Second revision date: March 30, 2004. Acceptance date: April 15, 2004.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Byung-Jo Kim, MD, PhD, Assistant Professor, Department of Neurology, Korea University College of Medicine, #126-1, Anamdong 5Ga, Seongbuk-Gu, Seoul, 136-705, Republic of Korea; E-mail: nukbj@korea.ac.kr

Spine reconstructive surgery has improved symptoms in carefully selected patients with incapacitating discogenic pain.[1] The goal of lumbar spine reconstruction is 3-fold: (1) to decompress neural compression; (2) to remove and reconstruct the affected intervertebral disc; and (3) restore instability and alignment of a pathologic motion segment. By removing the painful intervertebral disc and eliminating motion of anterior, middle, and posterior spinal structures, discogenic pain can be either reduced or eliminated from the pathologic segment.

The surgical treatment for patients with unremitting discogenic low back pain (LBP) has traditionally been spinal fusion, which is the most important part of recon-

structive surgery. The results of spinal fusion for chronic discogenic LBP are variable and have been controversial in terms of efficacy and type of fusion.[2] Reported outcomes for lumbar fusion in patients with chronic discogenic LBP have varied substantially.[3–7] In addition, there are significant morbidity and potential complications associated with these invasive procedures. Although some outcome studies have shown good clinical success rates (80% to 86%, assessed *via* patient satisfaction survey), others have reported less promising results.[8] Some investigators have even recommended only conservative approaches for treating discogenic LBP.[9]

Precise identification of pain generator(s), while difficult to quantify, is a key factor impacting clinical outcome. Other factors that putatively affect fusion outcome include smoking, age, psychological issues, and osteoporosis.[10–13] In addition, multilevel fusions have been thought to influence negatively success rates. Although there is general agreement that construct length adversely affects rates of arthrodesis,[14] this supposition has largely been based on anecdotal evidence. To our knowledge, only 2 reports in the literature analyze the effect of the number of levels on lumbar fusion surgery outcome.[15,16] One of these reports addresses fusion rate rather than symptom-oriented clinical outcome.[16] Most studies related to lumbar fusion outcome examine fusion at 1−2 levels only.[17–19] In addition, published studies have generally relied on magnetic resonance imaging or conventional diskography for diagnosis[15,16,18,19] rather than the more vanguard techniques, such as pressure-controlled diskography, which offer greater precision.[20] The purpose of this study is to determine if there is a difference of clinical outcome using Short Form RAND 36-Item Health Survey (SF-36) between short and long segment fusions.

■ Materials and Methods

Subjects. All data were obtained retrospectively from a prospectively collected database. Patients with chronic discogenic LBP, consecutively referred from primary care physicians and other clinicians during a period from January 1994 through August 2000 were eligible to participate in the study. The surgeon provided all patients with oral and written information about the treatment. Preoperative data were reviewed by a second surgeon who also participated in the study. For a patient to be included in the study, both surgeons had to concur regarding surgery indications.

Surgical methods included circumferential reconstruction of the lumbar spine by either posterior approach or combined anterior and posterior approach. Surgeries included posterior decompression necessary to relieve documented regions of neural compression, combined with interbody arthrodesis at selective levels, augmented by posterior segmental spinal instrumentation and posterolateral arthrodesis. Reconstruction of an interbody space was performed using a titanium cage (biomechanical) with hydroxyapatite (biologic) inserts combined with autologous growth factor (biochemical). It has been previously reported that this material may play a role in initial bone healing and stronger bone fusion, although controversy still exists as to its overall efficacy.[21,22]

Interbody arthrodesis was accomplished *via* either posterior or anterior approach. Interbody levels were selected for arthrodesis based on the results of provocative pressure-controlled diskography. The 2 senior surgical authors (JJL, TAK) performed all surgeries. Levels included in the fusion were determined in all patients by preoperative *independent* correlative testing for back pain symptoms using provocative pressure-controlled diskography. Symptoms referable to lower extremity pain were evaluated by selective foraminal epidural injection. A comprehensive conservative care program, including medical treatment, physical therapy, and spinal injection procedures, had failed for all patients. Patients were divided into 2 groups: a short segment group, involving interbody fusion at 1 or 2 levels and a long segment group, containing patients undergoing interbody fusion at 3–5 levels. Operated levels included those from L1−L2 through L5−S1.

All patients were expected to complete a preoperative, aquatic conditioning program to improve general health status. This program educated patients regarding postoperative rehabilitation and affirmed the commitment of the patient to be an active participant in their recovery. Whenever possible, coexisting medical conditions were corrected or stabilized before surgery. The situation of all patients involved in workers' compensation litigation resolved before the 2-year follow-up.

Data sources included medical records, radiographic review, and preoperative and postoperative questionnaires with SF-36 scores. A person blinded from patient information performed data entry *via* a network database.

Inclusion Criteria

Inclusion criteria were:

- Patients of both sexes, age 18–65 years
- Pain of at least 6 months duration failing conservative treatment
- Back pain more pronounced than leg pain
- Positive pressure-controlled diskography using concordant pain provocation criteria (pain score more than 6/10, pressure less than 50 psi above opening pressure).

Exclusion criteria were:

- History of major psychopathology
- Spinal fracture, infection, or neoplasm
- Inability to comply with the preoperative aquatic conditioning program

Outcome Measures

Initial Visit Questionnaire. Data were collected *via* initial visit questionnaire. Patients answered questions concerning spine symptoms, demographics, work, and the standardized SF-36 form. The majority of patients reported that they completed the entire questionnaire in less than 20 minutes. The treating physician completed questions about spine symptoms, diagnostic testing, specific spinal diagnosis, comorbidity, and treatment plans.

SF-36 Health Status Questionnaire. The functional status of patients was measured by SF-36 at preoperative baseline, and 1 and 2 years after surgery. Physical functional status captures how patients perceive their level of physical morbidity and its effects on their everyday lives. The 8 scales of the SF-36 are general health perceptions (GH), physical function (PF), general mental health (MH), role function as limited by physical

Table 1. Clinical Outcomes of Short and Long Segment Groups

SF-36 Variable	Short Segment Group			Long Segment Group		
	Preop	Postop 1-Yr	Postop 2-Yr	Preop	Postop 1-Yr	Postop 2-Yr
PF	25.65 (19.84)	51.41* (25.45)	57.28* (27.01)	26.13 (18.45)	42.98* (25.23)	45.11* (26.53)
RP	3.25 (13.13)	29.50* (40.73)	40.63* (43.46)	2.08 (8.35)	29.17* (40.69)	36.93* (42.62)
BP	21.20 (14.18)	48.02* (24.75)	48.45* (28.00)	17.93 (16.31)	43.20* (27.24)	41.67* (24.63)
GH	65.45 (17.54)	65.61 (21.17)	64.30 (22.87)	67.72 (17.82)	68.70 (22.73)	67.93 (21.07)
VT	36.30 (21.68)	49.85* (24.38)	50.00† (24.19)	41.29 (20.90)	53.23* (25.94)	51.85† (24.86)
SF	31.88 (20.67)	56.92* (28.63)	61.94* (32.11)	32.67 (25.93)	52.26* (33.36)	58.22* (32.17)
RE	46.34 (43.65)	63.33* (43.04)	67.91‡ (41.59)	43.88 (45.72)	57.22‡ (46.38)	52.24 (44.57)
MH	60.28 (23.35)	67.28* (22.72)	68.64 (21.72)	64.77 (22.22)	69.42 (25.17)	67.22 (22.34)
PCS	26.49 (7.31)	35.88* (10.28)	37.64* (11.13)	25.92 (6.05)	34.23* (10.09)	35.68* (10.38)
MCS	43.53 (12.96)	47.53* (12.97)	48.15 (12.46)	45.08 (12.42)	48.24‡ (13.36)	46.89 (12.98)

Data presented by mean (SD).
* $P < 0.005$; † $P < 0.01$; ‡ $P < 0.05$. All P values computed *via* paired t test.
PF = physical functioning; RP = role physical; BP = bodily pain; GH = general health; VT = vitality; SF = social functioning; RE = role emotional; MH = mental health.

problems (RP), role function as limited by emotional problems (RE), bodily pain (BP), vitality (VT), and social function (SF). These data were then used to compute Physical Component Summary (PCS) and Mental Component Summary (MCS) scores using the equation provided by the Medical Outcomes Trust.[23]

Statistical Analysis

All demographic data for the short and long segment groups were compared using the χ^2 test, except age data, which was analyzed *via* independent sample t-test. To compare the preoperative SF-36 scores for each group, an independent sample t-test was used. Preoperative and postoperative SF-36 scores in each group were compared using a paired t-test. Analyses of group differences in SF-36 score improvement were performed using an independent sample t-test. We compared the 1-year improvement in patients completing the 2-year follow-up with that of those completing only a 1-year follow-up by independent t-test. All data analysis was performed using SPSS software (Version 10, SPSS, Inc., Chicago, IL) by a researcher independent from the surgeons.

■ Results

Sample Characteristics

A total of 267 patients underwent circumferential reconstructive surgery. Of the 267 patients, 43 were eliminated due to incomplete SF- 36 questionnaires. Of the 224 patients who completed the SF-36 questionnaire without missing items (142 patients in the short segment group and 82 in the long segment group), 100 patients with short segment fusion (1.53 levels for mean; 47 at a single level and 53 at 2 levels), and 68 with long segment fusion (3.46 levels for mean; 44 at 3 levels, 17 at 4 levels, and 7 at 5 levels) completed the 1-year follow-up. Two-year follow-up was available for 81 patients in the short segment group (1.25 levels for mean; 37 at a single level and 44 at 2 levels), and 49 in the long segment group (3.39 levels for mean; 33 at 3 levels, 13 at 4 levels, and 3 at 5 levels). In the short segment group, interbody arthrodesis was undertaken with an approach through a single posterior incision. In the long segment group, 4-level and 5-level surgeries were performed with a combined anterior and posterior approach. Three-level surgeries used a single posterior approach.

Demographics

Men comprised 58.2% of the short segment group and 57.6% of the long segment group. Mean age was 41.0 years in the short segment group and 47.6 years in the long segment group. Smokers comprised 25.96% of patients in the short segment group and 26.6% of the long segment group. The percentage of patients in workers' compensation and litigation was 54.6% and 72.4% for the short segment group, and 53.0% and 72.1% for the long segment group, respectively. There were no significant differences between the 2 groups in gender, smoking habits, workers' compensation, or litigation ($P > 0.05$). No significant perioperative complications occurred in either group. There were no deaths in connection with surgery. In the short segment group, there was a 4.2% (6 of 142) complication rate, which included 1 dural laceration and 2 postoperative wound infections. In the long segment group, there was a 10.9% (9 of 82) complication rate, which included 5 dural lacerations, 1 temporary neural deficit, and 1 postoperative wound infection. All complications resolved with no obvious long-term sequelae. Pseudarthrosis occurred in 4 patients in the short segment group (2.8% for mean; 0 at a single level and 4 at 2 levels), and 19 in the long segment group (23.2% for mean; 13 at 3 levels, 3 at 4 levels, and 3 at 5 levels). Reoperation was performed for 2 patients in the short segment group (1.4% for mean; 0 at a single level and 2 at 2 levels), and 12 in the long segment group (14.6% for mean; 8 at 3 levels, 2 at 4 levels, and 2 at 5 levels).

Comparison of Outcome Between the 2 Groups

Mean preoperative PCS and MCS scores for the short and long segment groups did not differ significantly. Mean preoperative BP was lower in the long segment group ($P = 0.046$). Other scales were similar (Table 1).

For the short segment group, the postoperative 1-year mean PCS and MCS scores showed significant improve-

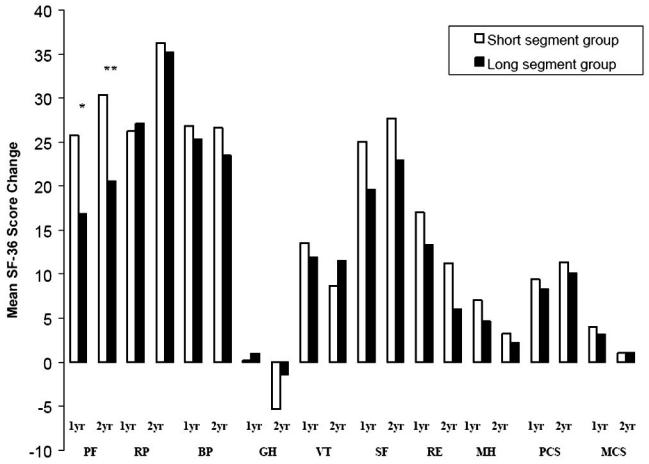

Figure 1. Changes in each mean SF-36 scale scores from preoperative to 1-year, and from preoperative to 2-year follow-up scores in the short segment and long segment groups (*$P < 0.05$, **$P < 0.1$). 1 yr, 1-year follow-up; 2 yr, 2-year follow-up; PF, physical functioning; RP, role physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role emotional; MH, mental health.

ment, increasing to 35.88 ($P < 0.001$) and 47.53 ($P = 0.002$, respectively). Scales other than GH showed significantly improved 1-year follow-up scores ($P \leq 0.001$). The 2-year follow-up scores showed significant improvement ($P < 0.001$ for PF, RP, BP, SF, and PCS). The VT and RE scores also improved, showing $P = 0.005$ and $P < 0.05$, respectively.

For the long segment group, the postoperative 1-year mean PCS score improved, increasing from 25.92 to 34.23 ($P < 0.001$). The MCS score also improved, increasing from 45.08 to 48.24 ($P < 0.05$). In addition, PF, RP, BP, and SF significantly improved ($P < 0.001$). The VT and RE scores also showed improvement ($P = 0.002$ and $P < 0.05$, respectively). Two-year follow-up PCS, PF, RP, BP, and SF scores also improved ($P < 0.001$). The VT score improved as well ($P < 0.01$) (Table 1).

Differences between preoperative and postoperative PCS and MCS scores suggest that the improvement in the long segment and short segment groups was comparable. Whole variances did not show significant differences ($P > 0.05$), except in the PF score, which was significantly smaller for the long segment group (16.85 ± 27.45 and 20.54 ± 27.81 at 1 and 2 years, respectively; $P = 0.048$) than for the short segment group (25.76 ± 27.81 and 30.37 ± 29.59 at 1 and 2 years, respectively; $P = 0.068$) (Figure 1).

Effect of Incomplete 2-Year Follow-up Data

To support the validity of results, we compared the 1-year follow-up data of patients completing both 1 and 2-year follow-up, with the 1-year follow-up data of patients completing the 1-year follow-up only. No significant differences were found ($P > 0.1$) (Table 2).

■ Discussion

Lumbar reconstructive surgery with fusion is commonly performed to treat segmental instability and pain caused by various spinal disorders. The most common indication for reconstructive surgery in patients with painful disc disruption is disabling discogenic LBP.[1] Reported pain is typically mechanical and intolerant to axial activities such as, standing and sitting. The pain often worsens with activity. Numerous studies have reported surgical results for patients with internal disc disruption, discogenic pain, or painful degenerative disc disease.[3–7] Successful clinical outcomes in these studies range from 45% to 90%.

Consideration of fusion procedures for patients with discogenic pain remains controversial. Careful patient selection is critical. In addition, successful fusion with current techniques may be problematic with longer fusion segments. Narayan *et al*[16] investigated the correlation of the number of fusion levels with successful arthrodesis rates in 457 patients who underwent

Table 2. Comparison of 1-year improvement in patients completing 2-year follow-up and patients completing only 1-year follow-up

	Short Segment Group					Long Segment Group				
SF-36 Items	2-Year Follow-Up	No 2-Year Follow-Up	*t*	df	Sig.* (2-tailed)	2-Year Follow-Up	No 2-Year Follow-Up	*t*	df	Sig.* (2-tailed)
PF	16.67 (26.45)	17.25 (30.15)	−0.08	60	0.938	26.22 (29.36)	24.85 (24.91)	0.23	98	0.817
RP	32.50 (42.44)	16.25 (39.13)	1.43	58	0.157	29.17 (41.97)	20.59 (40.58)	0.98	98	0.330
BP	25.39 (31.57)	25.00 (30.25)	0.05	59	0.964	29.42 (28.59)	21.76 (26.53)	1.30	98	0.197
GH	1.23 (24.76)	0.47 (28.57)	0.11	60	0.915	−0.64 (30.12)	1.72 (23.54)	−0.40	98	0.692
VT	12.5 (25.03)	15.59 (25.07)	−0.58	98	0.560	13.69 (29.05)	8.25 (29.57)	0.68	37	0.500
SF	17.30 (38.74)	24.40 (39.98)	−0.67	60	0.507	26.36 (34.72)	22.49 (28.79)	0.56	98	0.577
RE	12.50 (36.75)	15.02 (59.69)	−0.20	58	0.841	14.14 (45.70)	22.54 (55.53)	−0.76	57	0.451
MH	5.43 (18.57)	3.00 (24.10)	0.44	60	0.664	5.27 (24.60)	10.35 (20.16)	−1.04	98	0.302
PCS	8.92 (11.97)	7.10 (13.60)	0.53	58	0.599	10.39 (13.56)	7.45 (11.87)	1.07	98	0.287
MCS	3.11 (10.17)	3.28 (14.97)	−0.05	58	0.960	2.94 (12.23)	6.05 (12.95)	−1.18	98	0.242

Data presented by mean (SD). The 1-year follow-up score was subtracted from the preoperative score.
* Statistical analysis performed by independent t-test.
PF = physical functioning; RP = role physical; BP = bodily pain; GH = general health; VT = vitality; SF = social functioning; RE = role emotional; MH = mental health; df = degree of freedom; sig = significance.

posterolateral intertransverse lumbar fusion with pedicle screw instrumentation. Using a criterion for successful fusion based on the radiographic demonstration of a bilateral contiguous osseous bridge over the transverse processes and absence of movement on dynamic radiograph films, the success rate in patients with single-level degenerative disc disease was 91%. Fusion rates declined steeply as a function of each additional motion segment in the translational instability group. These data suggested that arthrodesis success rates appear to be related to the number of segments included in the fusion. However, this study did not address the outcome of clinical symptoms and quality of life. Glassman *et al*[15] reported more deterioration in the general health of a 3-level fusion group in an outcome study of 30 patients who underwent decompression and instrumented fusion using a pedicle screw/rod construct and autologous iliac crest bone graft. Greater improvement in SF and MH was seen in 1-level fusions, but these results did not reach statistical significance due to the predominance of 2-level fusions in the SF-36 data. It is noteworthy that these studies did not use precise pressure-controlled diskography as a diagnostic tool.

In our study, pseudarthrosis and reoperation rates were significantly higher in the long segment group. Successful arthrodesis is the fundamental surgical goal in lumbar reconstructive surgery. However, successful fusion does not ensure clinical success.[1] When clinical outcome related to quality-of-life was compared, the short and long segment groups afforded similar results. Mean PF score increment was significantly less for the long segment group (16.9 and 19.0 at 1 and 2-year follow-ups, respectively) than for the short segment group (25.8 and 31.6 at 1 and 2-year follow-ups, respectively). Other SF-36 variances did not show significant differences between the 2 groups. The PCS score, which summarizes the individual scales with an emphasis on physical behavior, showed similar improvements. The PCS score increments for the short segment group were 9.4 (from 26.5 to 35.9) and 11.1 (from 26.5 to 37.6) at 1 and 2-year follow-up, respectively. The PCS score increments for the long segment group were 8.3 (from 25.9 to 34.2), and 9.8 (from 25.9 to 35.7) at 1 and 2-year follow-up, respectively. This result is comparable to that of other lumbar fusion outcome studies. In an efficacy study comparing the use of recombinant human bone morphogenous protein-2 (rhBMP-2) with the use of autogenous iliac crest bone graft after lumbar interbody fusion at a single level, mean PCS scores increased approximately 15 (from 30 to 45) and 10, respectively.[18]

A critical factor for successful clinical outcome in lumbar fusion surgery is patient selection. We used strict diagnostic criteria for identification of pain generators. Although the diagnostic power of diskography remains controversial,[24] the skilled use of pressure-controlled diskography with strict criteria for positive findings can increase diagnosis specificity. Previous studies have supported the importance of accurate diagnosis in clinical outcome.[20] We selected fusion levels *via* pressure-controlled diskography using concordant pain provocation criteria of pressure less than 50 psi above opening pressure with Grade 3 anular tears.[20] We also confirmed negative discs above and below fused levels. Although this study was not specifically designed to evaluate the predictive power of pressure-controlled diskography, a thorough preoperative assessment of discogenic pain helped ensure that all potential painful segments were included. This approach may have been valuable for the treatment of patients in the long segment group, in which the precise analysis of a higher number of levels is required.

The type of surgery performed may have also affected the outcome of longer segment fusions. The use of a combined procedure including both posterior and anterior reconstruction with posterior segmental spinal instrumentation and posterolateral arthrodesis may increase fusion success rates, affording improvement in subjective outcome. Furthermore, a well-designed preoperative conditioning program, including aquatic therapy, and stabilization of medical comorbidity for improving general health may have been a factor for the clinical outcome related to quality-of-life in each group.

One of limitations of this study is incomplete 2-year follow-up data. Although patients enrolled in the study were tracked, some were unavailable after 2 years to complete the final follow-up. However, these missing data correspond to a random loss of data points and do not systematically impact study outcome. Comparison of 1-year follow-up data of patients completing both 1 and 2-year follow-up, with 1-year follow-up data of patients completing only the 1-year follow-up, showed no significant differences (Table 2). Because the 2-year outcomes show an improvement trend similar to that of 1-year outcomes, we can conclude that the effect of an incomplete 2-year outcome data set is negligible. Another potential shortcoming is the extended interval during which the study data gathered (*i.e.*, 1994 to 2000). Although the authors attempted to ensure that all patients underwent identical study protocols, slight variations in technique or procedure may have occurred due to the long study interval. In addition, this study did not exclude the influence of postoperative pain treatment on the clinical outcome during the 2-year follow-up. Due to the retrospective nature of this study, patients were not limited in the manner in which they sought to control LBP during follow-up.

The SF-36 questionnaire has been used extensively to assess general self-reported quality of life in a variety of populations. Typically, general measurements such as the SF-36 are combined with condition-specific instruments, such as the Oswestry Disability Index. However, the use of both types increases respondent burden, redundancy, data collection, and analysis burden.[25] The use of PCS for spine research has been recommended as a dependent measure instead of the 8 SF-36 subscales be-

cause the number of statistical analyses may be reduced without missing clinical differences.[23,26]

In contrast to previous spine surgery outcome studies based on SF-36 data, our results clearly show significant clinical improvement. The relatively few papers reporting the clinical outcome of lumbar fusion are based on the analysis of PF, BP, and PCS variances, and often do not include a detailed analysis of the complete set of SF-36 variances.[15,18,27] In our study, MH and MCS did not show significant change, while the GH after the procedure was adversely affected. It is possible that these results may be related to various factors impacting patient emotional status (*e.g.*, workers' compensation, litigation, and psychological susceptibility). Further study is needed to clarify the correlation between the emotional-behavior-focused SF-36 variances, psychological scales, and secondary gain.

In an SF-36-based outcome study of total joint arthroplasty, a well-accepted and successful treatment modality for end-stage arthritis,[28] PCS score increments of 10 and 12 were reported.[28] Similar PCS score improvements obtained in this study (11.35 and 10.06 for the short and long segment groups during 2-year follow-up, respectively) support the use of lumbar fusion in carefully selected patients as an effective, reliable treatment modality for discogenic pain.

These data have shown results for the long segment group in general. However, for 5-level fusions, mean PCS and MCS score improvements at 1-year follow-up (n = 7) were not significant (24.57 to 29.73 and 41.46 to 38.86, respectively). Two-year follow-up results could not be compared due to the small sample size. Pseudarthrosis and reoperation rates were significantly higher in the long segment group than the short segment group. For cases in which pseudarthrosis occurred in the long segment group, mean PCS score improvement at 1-year follow-up was not significant (27.1 to 32.43). Although there have been study results that showed no significant correlation between successful fusion and overall clinical results for the patient, or improvement in symptoms,[1,12] the occurrence of higher pseudarthrosis rates may remain controversial for long segment fusion in patients with chronic discogenic LBP.

■ Conclusion

The effect of the number of levels on combined lumbar reconstructive surgery clinical outcome has been investigated using provocative pressure-controlled diskography as the primary diagnostic tool for precisely determining surgical levels. Mean PF score increment was significantly less for the long segment (3–5 level) than the short segment (1–2 level) group. There were no statistical differences in the other variances of the SF-36 scales, including PCS score. Although some surgical results (*e.g.*, nonunion and reoperation rates) were significantly higher in the long segment group, the number of levels in lumbar fusion surgery may not significantly impact clinical outcome measured by SF-36 when strict patient selection criteria, precise identification of pain generators, and a preoperative conditioning program for improving general patient health status are used. The clinical improvement for lumbar reconstruction measured by SF-36 was comparable to treatment outcomes of other chronic diseases and well-established surgical treatments.

■ Key Points:

- The number of intervertebral disc levels fused during lumbar fusion surgery may not significantly impact the clinical outcome measured by SF-36.
- Spinal reconstructive surgery, combined with a well-designed and thorough preoperative program including appropriate patient selection, may offer significant improvements during the postoperative quality of life.

Acknowledgment

The authors thank Todd M. Billeci for editing the manuscript.

References

1. Slosar PJ. Indications and outcomes of reconstructive surgery in chronic pain of spinal origin. *Spine* 2002;27:2555–62.
2. Biyani A, Andersson GB, Chaudhary H, et al. Intradiscal electrothermal therapy: A treatment option in patients with internal disc disruption. *Spine* 2003;28:S8–S14.
3. Barrick WT, Schofferman JA, Reynolds JB, et al. Anterior lumbar fusion improves discogenic pain at levels of prior posterolateral fusion. *Spine* 2000; 25:853–7.
4. Buttermann GR, Garvey TA, Hunt AF, et al. Lumbar fusion results related to diagnosis. *Spine* 1998;23:116–27.
5. Cloward RB. The treatment of ruptured lumbar intervertebral discs by vertebral body fusion. I. Indications, operative technique, after care. *J Neurosurg* 1953;10:154–68.
6. Fraser RD. Interbody, posterior, and combined lumbar fusions. *Spine* 1995; 20(suppl):167S–77S.
7. Zdeblick TA. A prospective, randomized study of lumbar fusion. Preliminary results. *Spine* 1993;18:983–91.
8. Knox BD, Chapman TM. Anterior lumbar interbody fusion for discogram concordant pain. *J Spinal Disord* 1993;6:242–4.
9. Smith SE, Darden BV, Rhyne AL, et al. Outcome of unoperated discogram-positive low back pain. *Spine* 1995;20:1997–2000.
10. Brown CW, Orme TJ, Richardson HD. The rate of pseudarthrosis (surgical nonunion) in patients who are smokers and patients who are nonsmokers: A comparison study. *Spine* 1986;11:942–3.
11. Schnee CL, Freese A, Ansell LV. Outcome analysis for adults with spondylolisthesis treated with posterolateral fusion and transpedicular screw fixation. *J Neurosurg* 1997;86:56–63.
12. Goldner JL. Symposium: The role of spine fusion for low-back pain. International Society for the Study of the Lumbar Spine, New Orleans, LA, May 27, 1980. *Spine* 1981;6:293–303.
13. Polatin PB, Kinney RK, Gatchel RJ, et al. Psychiatric illness and chronic low-back pain. The mind and the spine–Which goes first? *Spine* 1993;18: 66–71.
14. Hanley EN, Phillips ED, Kostuik KA. *The Adult Spine*. New York, NY: Raven Press; 1991.
15. Glassman SD, Minkow RE, Dimar JR, et al. Effect of prior lumbar discectomy on outcome of lumbar fusion: A prospective analysis using the SF-36 measure. *J Spinal Disord* 1998;11:383–8.
16. Narayan P, Haid RW, Subach BR, et al. Effect of spinal disease on successful arthrodesis in lumbar pedicle screw fixation. *J Neurosurg* 2002;97(suppl): 277–80.
17. Fritzell P, Hagg O, Wessberg P, et al. Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: A multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001;26:2521–32.

18. Burkus JK, Transfeldt EE, Kitchel SH, et al. Clinical and radiographic outcomes of anterior lumbar interbody fusion using recombinant human bone morphogenetic protein-2. *Spine* 2002;27:2396–408.
19. Ivar Brox J, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913–21.
20. Derby R, Howard MW, Grant JM, et al. The ability of pressure-controlled discography to predict surgical and nonsurgical outcomes. *Spine* 1999;24:364–71.
21. Wong CT, Lu WW, Chan WK, et al. In vivo cancellous bone remodeling on a strontium-containing hydroxyapatite (sr-HA) bioactive cement. *J Biomed Mater Res* 2004;68A:513–21.
22. Lowery GL, Kulkarni S, Pennisi AE. Use of autologous growth factors in lumbar spinal fusion. *Bone* 1999;25(suppl):47S–50S.
23. Ware J, Kosinski M, Keller S. *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston, MA: The Health Institute; 1994.
24. Carragee EJ. Prevalence and clinical features of internal disk disruption in patients with chronic low back pain. *Spine* 1996;21:776–7.
25. Walsh TL, Hanscom B, Lurie JD, et al. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. *Spine* 2003;28:607–15.
26. Fanuele JC, Birkmeyer NJ, Abdu WA, et al. The impact of spinal problems on the health status of patients: Have we underestimated the effect? *Spine* 2000;25:1509–14.
27. Debnath UK, Freeman BJ, Gregory P, et al. Clinical outcome and return to sport after the surgical treatment of spondylolysis in young athletes. *J Bone Joint Surg Br* 2003;85:244–9.
28. Mahomed NN, Liang MH, Cook EF, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol* 2002;29:1273–9.