SPINE Volume 26, Number 7, pp 738–747

Outcomes of Posterolateral Lumbar Fusion in Utah Patients Receiving Workers' Compensation

A Retrospective Cohort Study

M. Scott DeBerard, PhD,* Kevin S. Masters, PhD,* Alan L. Colledge, MD,†
Rand L. Schleusener, MD,‡ and John D. Schlegel, MD‡

Study Design. A retrospective cohort study consisting of a medical record review and a follow-up telephone survey of patients with lumbar fusion, at least 2 years after their surgery, was performed.

Objective. To identify presurgical correlates and long-term outcomes from posterolateral lumbar fusion in Utah patients receiving workers' compensation.

Summary of Background Data. Lumbar fusion has been criticized for its highly variable outcomes, and compensated workers are at particular risk for poor outcomes. Evidence suggests that presurgical psychosocial factors may be important modifiers of back pain reporting and back surgery outcomes.

Methods. The patients in this study were 185 compensated workers in Utah who underwent posterolateral lumbar fusion. Patient medical records were independently reviewed, and medical and sociodemographic variables were coded. A telephone outcome survey was completed with 130 patients (70%) an average of 4.6 years after their surgery.

Results. Reported solid fusion, reoperation, and disability rates for the follow-up sample were 74%, 24%, and 25%, respectively. As reported by the patients, 41% experienced no change or a worsened quality of life. Mean scores from the Roland and Morris Back Pain Disability Questionnaire, the Stauffer-Coventry-Index, and the Short-Form 20 Multidimensional Health Survey indicate that many patients experienced postsurgical dysfunction. Presurgical predictors of outcomes were number of prior low back operations, income at time of injury, age, litigation, and depression.

Conclusions. Outcomes of posterolateral lumbar fusion among compensated workers in Utah are inconsistent. Outcomes can be predicted by presurgical sociodemographic variables. Screening for such presurgical risk factors may be important for prudent surgical decisions and rehabilitation planning. [Key words: arthrodesis, lumbar fusion, patient outcomes] **Spine 2001;26:738–747**

Low back pain is a significant economic and psychosocial problem affecting 15% to 20% of the working-age population annually in the United States.[2,7] Approximately 5% of the patients with acute low back pain do not respond to conservative therapies (*e.g.*, physical therapy, patient education, over-the-counter pain relievers, light exercise) and experience chronic and disabling pain.[9] Many of these patients, particularly injured workers, turn to surgical intervention as a next treatment option.

Approximately 192,000 patients with chronic, intractable back pain undergo spinal fusion each year.[15] It is assumed that surgical fusion can resolve instability of vertebral bodies and consequently reduce pain and disability. Whereas a number of lumbar fusion techniques are used currently, the focus of this study was on posterolateral fusion with and without instrumentation.

Posterolateral lumbar fusion, particularly with instrumentation, has been criticized for producing inconsistent results.[5,27] For example, Turner et al[27] conducted a review of 47 published spinal fusion studies (1966–1991) and found that the percentage of patients with satisfactory outcomes varied from 16% to 95%, averaging 68%. In a recent study of compensated lumbar fusion outcomes, most patients reported that their back pain was worse (67.7%) and their overall quality of life was worse or no better (55.8%) than before surgery, as measured at a 2-year follow-up assessment.[5] Many possible surgically related explanations for poor lumbar fusion outcomes have been offered including soft tissue injury secondary to a protracted operative period; rigorous dissection, movement, or breaking of implanted hardware; increased number of vertebral levels fused; and pseudoarthrosis. Presurgical sociodemographic factors including older age, litigation, increased number of prior low back surgeries, increased time of work disability, low income, compensation, and psychological status also have been implicated prognostically in poor back surgery outcomes.[1,3–5,13,14,20,24,25,27–31] Other potential reasons for inconsistent lumbar fusion outcomes include a lack of consensus regarding indications for lumbar fusion,[10,11] variations in patient selection protocols and postsurgical rehabilitation,[12] and a lack of a unified criteria for assessing surgical and clinical outcomes in lumbar fusion research studies.[10]

Given the multiplicity of factors that may influence recovery from lumbar fusion, it is useful to identify a model that can integrate relevant variables and guide research protocols. The authors believe that recovery processes relative to lumbar fusion are best conceptualized from a biopsychosocial model of illness.[26] This

From the *Department of Psychology, Utah State University, Logan, Utah, the †Labor Commission of Utah, Salt Lake City, Utah, and the ‡Department of Orthopedics, University of Utah Medical School, Salt Lake City, Utah.
Supported from grants provided by the University of Utah Medical School and the Labor Commission of Utah.
Acknowledgment date: October 6, 1999.
First revision date: February 28, 2000.
Acceptance date: May 15, 2000.
Device status category: 11.
Conflict of interest category: 12.
The results presented in this article should not be construed as an endorsement or indictment of lumbar fusion techniques for injured workers by the authors, Utah State University, the Labor Commission of Utah, or the University of Utah.

model posits that biologic, social, and psychological factors all are implicated in any given state of health or illness. Alternatively, the biomedical approach suggests that pathologic physiologic processes are the sole determinants of illness, and therefore that psychological and social factors are independent of disease and healing processes.[26] If the biopsychosocial model is a valid approach to conceptualizing recovery from lumbar fusion, it is expected that social and psychological factors before surgery as well as biologic factors should be predictive of outcomes.

The basic purpose of this study was to examine presurgical correlates and postsurgical outcomes of posterolateral lumbar fusion in injured workers from Utah. An additional purpose was to determine whether presurgical variables reflective of the biopsychosocial model were useful for predicting outcomes of posterolateral lumbar fusion.

■ Methods

Study Design. This retrospective cohort study consisted of coding presurgical information documented in patient medical records and conducting a telephone outcome survey with patients at least 2 years after their surgeries. The institutional review boards from the University of Utah Medical School and the Utah State University approved this study, and the Workers' Compensation Fund of Utah (WCFU) provided permission to access patient medical records and contact patients for follow-up evaluation.

Patient Sample. Patients were eligible for inclusion in this study if they had undergone posterolateral lumbar fusion between 1990 and 1995, had no presurgical diagnosis of vertebral fracture, and were at least 2 years beyond surgery at the time of follow-up assessment. Patients were identified *via* current procedural terminology (CPT) codes in WCFU databases. For each of the 185 eligible patients, lumbar fusion resulted from a verified workplace low back injury and the WCFU system covered medical and rehabilitation expenses and lost wages. Workers covered by federal workers' compensation systems and self-insured employers were excluded because of inaccessible data. The WCFU insures approximately 55% of Utah workers.

Medical Record Information. Presurgical medical record data were gathered *via* independent and standardized review of medical chart information and review of WCFU computer databases. Data abstractors were trained graduate students uninvolved with the treatment of study patients, who routinely met with authors to review and resolve coding problems. The sociodemographic variables coded for this study included age at the time of injury, household income at the time of injury, time delay from injury to surgery, smoking status at the time of surgery, current depressive disorder (*Diagnostic and Statistical Manual of Mental Disorders* [DSM-IV] major depressive disorder, dysthymia, or adjustment disorder with depressed mood) documented in preoperative medical records, litigation as defined by patient private lawyer involvement in the compensation case at the time of surgery, number of prior low back operations, number of vertebral levels to be fused, and a presurgical diagnostic severity score based on independent review of presurgical radiology reports by three of the authors (A.L.C., R.L.S., J.D.S.). A copy of the presurgical diagnostic severity measure is included in the Appendix. Preoperative diagnosis and surgical information (*i.e.,* instrumentation, hospital) also were coded.

Assessment of Arthrodesis Rates. Arthrodesis was assessed by a review of the operating surgeon's postoperative findings recorded in the patient medical chart. For all patients, the operating surgeons routinely reviewed radiographs and documented the progression of bone consolidation. The operating surgeon's final determination of arthrodesis, based on radiographic imaging findings, was coded as a dichotomous variable (yes or no). The radiologists' reports of patient radiographs were found in the medical chart. These were reviewed independently by an orthopedic surgeon (R.L.S.) and occupational physician (A.L.C.) to verify the surgeon's accurate interpretation. The reviewers had no involvement in patient treatment. The operating surgeon's final opinion of arthrodesis typically was determined and documented in the chart at the time of the patient's maximal medical improvement, an average of 1.5 years after surgery.

Patient Outcome Survey. An objective outcome instrument, including an interview script for telephone interviewers, was created. This instrument included three patient satisfaction items drawn from a recent study of compensated fusion patients: quality of life improvement resulting from lumbar fusion, perceived back or leg pain improvement, and satisfaction with back condition at the time of follow-up evaluation.[5] Disability status at follow-up evaluation was assessed by asking subjects whether they currently were receiving total disability benefits for their back condition.

The Roland and Morris Back Pain Disability Questionnaire (DQ), a 24-item self-report measure designed to evaluate dysfunction associated with low back pain, also was used.[18,19] The Stauffer-Coventry Index was selected as a clinical surgical outcome measure.[21] This index, a widely used measure for assessing low back surgical outcomes,[21,28] was used as an outcome measure in at least two relatively recent reviews.[25,27] Designed for administration after surgery, it consists of four multiple-choice self-report questions regarding pain reduction, return to work, limitations of physical activities, and medication usage. On the basis of their lowest rated category, patients may be assigned to one of three possible clinical outcome groups: 1) good (76% to 100% relief in leg and back pain, return to previous work status, minimal or no restriction of physical activities, occasional mild analgesics or no analgesics), 2) fair (26% to 75% relief of leg and back pain, return to lighter work, moderate restrictions of physical activities, regular use of nonnarcotic analgesics, and 3) poor (0% to 25% relief of leg and back pain, no return to work after surgery, severe restrictions of physical activities, occasional or regular use of narcotic analgesics).

The Short-Form-20 Multidimensional Health Survey (SF-20),[22,23] a 20-item general health survey, was used to assess six general dimensions of health: 1) physical functioning, 2) role functioning, 3) social functioning, 4) mental health, 5) current overall health rating, and 6) pain.

Procedures For Outcome Survey. The initial contact with patients occurred *via* a patient letter sent to their most recent address identified in the medical records. This letter, sent to all the patients who met the inclusion criteria, explained the re-

search purposes of the study, procedures and assurances of confidentiality, and a request for their participation. A self-addressed, stamped card was included so patients could inform researchers of any changes in address or telephone numbers and their agreement to participate in the study.

The patients who mailed their cards back were contacted immediately for the survey. The patients who did return their cards were sent as many as two additional contact letters requesting their participation. If they still did not respond, a telephone contact was attempted. The survey was completed at that time if the patient consented verbally. If a patient's telephone number was different from that listed in existing databases, an attempt was made to locate the patient through calling directory assistance or two Internet search programs: Lycos and People Find. The telephone outcome surveys were conducted by three graduate students who had no treatment contact with the patients. The interviewers were trained in basic interviewing skills and provided with a detailed written script to follow when conducting the survey. All the surveys were completed in one session approximately 20 to 75 minutes long. The possibility of winning a drawing of $500 was offered to patients as an incentive to participate.

■ Results

Presurgical Patients and Follow-Up Data

The study cohort included 185 patients who had received their surgeries between August 23, 1990 and April 21, 1995. The medical charts of 150 men (81.1%) and 35 women (18.9%) were reviewed. Of this sample, 96% were white, 3% were Hispanic, and 1% were Native American. Preoperative diagnoses for the 185 patients were as follows: disc herniation (68.5%), degenerative disc disease (50.8%), segmental instability (27.7%), spondylolisthesis (16.9%), spinal stenosis (20%), and pseudoarthrosis (3.1%). Multiple presurgical diagnoses were possible. All the patients had received posterolateral lumbar fusions, and instrumentation had been used in 83.7% of the patients. Pedicle screw fixation was used in 91% of instrumented cases. In all, 28 surgeons were involved in the 185 operations. The overall follow-up rate for the patient outcome survey was 70% (130/185). The average time to follow-up assessment was 4.6 years.

A multivariate analysis of variance (MANOVA) comparing respondents and nonrespondents on the nine presurgical characteristics was conducted to determine possible response biases. This analysis, shown in Table 1, determined that respondents were slightly older and slightly less likely to smoke than nonrespondents. Because these differences were very small (see η-squared values in Table 1), it was assumed that the results from the outcome survey were not differentially biased on the basis of assessed presurgical characteristics.

Fusion and Reoperation Rate

The reported arthrodesis rate for the total study sample was 73% (135/185) and 73.8% (96/130) for the follow-up sample. Arthrodesis rates did not vary among respondents, as compared with the rates among nonrespondents, adding support to the notion that the follow-up sample was not differentially biased in terms of surgical outcomes. The reoperation rate for the follow-up sample was 23.8% (31/130).

Table 1. Descriptive Statistics and Comparisons of Presurgical Medical and Sociodemographic Characteristics For Respondent *vs* Nonrespondent Patients

Presurgical Variable	Means or Proportion: Respondents (n = 130)	Means or Proportion: Nonrespondents (n = 55)	P	η^2
Age, yrs	38.7	34.5	0.004*	0.044
Household income at time of injury, $	420.7	369.8	0.097	0.015
Time delay from injury to surgery, days	427.6	294.9	0.072	0.018
Reported as smoker, %	43.8	60.0	0.045*	0.022
Presurgical diagnosis of depression, %	9.0	15.3	0.290	0.006
Lawyer involvement in compensation case, %	38.5	45.5	0.379	0.004
Prior low back operations, %				
No prior	50.8	58.2		
1 prior	38.5	27.3		
2 prior	9.2	7.3		
3 or more	1.5	7.3	0.867	0.000
Number of Levels Fused, %				
1 level	46.9	49.1		
2 level	48.5	49.1		
3 or more levels	4.6	1.8	0.589	0.002
Mean imaging severity	7.1	8.0	0.195	0.000

* $P < 0.05$.

Patient Satisfaction Items

Patient responses to the four patient satisfaction items are presented in Table 2.

Table 2. Patient Satisfaction Outcomes

Outcome Category	Frequency	%
Quality of life improvement resulting from lumbar fusion surgery		
Great improvement	37	28.5
Moderate improvement	28	21.5
Little improvement	12	9.2
No change	11	8.5
Little worse	7	5.4
Moderately worse	15	11.5
Much worse	20	15.4
Is your back or leg pain problem better than, worse than, or what you expected it to be at this point		
Much better	26	20.0
Somewhat better	25	19.2
What I expected	32	24.6
Somewhat worse	19	14.6
Much worse	28	21.5
Satisfaction with back condition as it is right now		
Extremely dissatisfied	24	18.5
Very dissatisfied	9	6.9
Somewhat dissatisfied	16	12.3
Neutral	26	20.0
Somewhat satisfied	30	23.1
Very satisfied	15	11.5
Extremely satisfied	10	7.7

Table 3. Stauffer-Coventry Index: Subscale Scores and Aggregate Ratings

Pain Relief			Employment Status			Physical Limitations			Medication Usage			Overall Index Rating*		
Category	Frequency	%	Category	Frequency	%	Category	Frequency	%	Category	Frequency	%	Category	Frequency	%
Good (76–100% improvement)	42	32.3	Good (return to previous work status)	33	25.4	Good (minimal or no restrictions)	22	16.9	Good (occasional or no use of mild analgesics)	57	43.8	Good	9	6.9
Fair (26–75% improvement)	56	43.1	Fair (return to lighter work)	66	50.8	Fair (moderate restrictions)	67	51.5	Fair (regular use on nonnarcotic analgesics)	39	30.0	Fair	56	43.1
Poor (0–25% improvement)	32	24.6	Poor (no return to work)	31	23.8	Poor (severe restrictions)	41	31.5	Poor (occasional or regular use of narcotic analgesics)	34	26.2	Poor	65	50.0

Note: Percentages based on follow-up n of 130 patients.
* Final classification based on lowest rated single category.

Disability Status

At follow-up assessment, 25% (32/130) of the follow-up cohort were totally and permanently disabled.

Roland and Morris Back Pain Disability Questionnaire

Patients obtained a mean score of 11.4 (SD = 6.96), with a minimum score of 0 and a maximum score of 24. On the basis of the recommendation from the original articles, a cutoff of 14 or more points was used to determine poor outcome.[18,19] Consequently, 43.1% of the follow-up group fell into the poor outcome range. This relatively high percentage of poor outcomes stands in contrast to the 15% found in the original standardization sample of patients with low back pain (nonsurgical),[18,19] suggesting that after surgical fusion, patients in this sample were more disabled than generally is typical of patients with low back pain.

Stauffer-Coventry Index

The four subscale values and aggregate ratings for the Stauffer-Coventry Index are presented in Table 3. The aggregate ratings show that only 6.9% of the patients achieved good outcomes, whereas 50% experienced poor results. However, if averages across subscales were used to determine total aggregate scores instead of the lowest rated single category, then the percentage of good outcomes would be substantially higher (29.6%). This difference is partly attributable to the fact that many patients with otherwise good outcomes in terms of return to work, pain relief, and medication usage still indicated at least moderate physical limitations after surgery, thus relegating them to the fair outcome category. The authors believe that the aggregate index may underrepresent the number of patients who had an overall good outcome from lumbar fusion. Nevertheless, the aggregate index was used in subsequent correlational analyses to maintain the reliability and validity associated with established normative scoring criteria.

Short-Form 20 Multidimensional Health Survey

Table 4 contains the follow-up sample means and standard deviations for the six SF-20 subscales along with normative sample means and standard deviations for medical outpatients, as provided by Stewart and Ware.[22,23] The follow-up means generally indicate a trend of overall poorer perceived health among the post-fusion sample. To further characterize these differences between the follow-up and normative samples, a standardized mean difference was calculated for each sub-

Table 4. Descriptive Statistics for the Short-Form 20 (SF-20) Multidimensional Health Survey Subscales

				Percentage in Poor Health‡		
SF-20 Subscale*	Follow-up Mean (SD)	Normative Mean (SD)†	Effect Size	Fusion Sample	Normative Patients†	General Population§
Physical functioning (6 items)	46.2 (29.6)	78.5 (30.8)	−1.05	91	45	22
Role functioning (2 items)	48.3 (42.9)	77.5 (38.3)	−0.76	66	28	12
Social functioning (1 item)	71.7 (34.1)	87.2 (23.6)	−0.66	27	9	NA
Mental health index (5 items)	62.7 (22.9)	72.6 (20.2)	−0.49	51	31	19
Current health perceptions (5 items)	53.7 (27.7)	63.0 (26.8)	−0.35	66	52	20
Pain severity (1 item)	57.2 (22.7)	31.4 (27.7)	0.93	67	29	NA

* Observed range of all scores was 0–100. A high score indicates better health except for pain, where a high score indicates more pain.
† The normative sample consisted of patients presenting to physicians, psychologists, and other mental health providers within health maintenance organizations (HMOs), multispecialty groups, and solo-fee-for-service groups.
‡ Poor health defined as physical and role functioning means one or more limitations; social functioning means limitations much of the time, mental health means the lowest 19% of the scores in the general population sample (score of 67 or lower; cutoff defined as close as possible to the bottom 20%); health perceptions means the lowest 20% of the scores in the general population sample (score of 70 or lower); pain means moderate, severe, or very severe pain.
§ General population refers to a random sample of patients selected for telephone administration of the SF-20.[22]
NA = not available.

Table 5. Logistic Regression Equation Predicting Disability Status, With Nine Presurgical Variables as Predictors

Variable	β	P	Exp (B)	95% CI
Age (5-year increments from age 25 years)	0.7823	0.0002	2.1866	(1.45–3.29)
Household income ($100 increments from $0.00)	−0.4038	0.0133	0.6678	(0.49–0.92)
Time delay (1-year increments from 0 years)	−0.1115	0.5130	0.8945	(0.64–1.25)
Smoking	−0.2262	0.6752	0.7976	(0.28–2.30)
Depression	1.2721	0.1200	3.5684	(0.72–17.74)
Litigation	1.5604	0.0046	4.7608	(1.60–13.99)
Prior low back operations	0.6782	0.0738	1.9703	(0.94–4.14)
Levels fused	1.0739	0.0220	2.9267	(1.17–7.34)
Imaging severity rating (4-point increments from 0)	−0.2478	0.3806	0.7805	(0.45–1.36)
Constant	−8.0155	0.0004		

Exp (B) = estimated odds ratio; CI = critical interval for exp (B).

scale. Overall, the current sample demonstrated significantly lower perceived health status than that of normative medical patients. This difference was evidenced by large (>0.8) to small (~0.2) effect sizes across a number of diverse health indexes.

Use of Presurgical Variables in Predicting Patient Outcomes

Tables 5 to 11 show the results of various analyses using the nine presurgical variables to predict patient outcomes. Table 5 depicts the results from a multiple logistic regression analysis of the presurgical patient variables on postsurgical disability status. Age, income, time delay from injury to fusion, and presurgical imaging severity were recoded from their original continuous form to a continuous equal-interval form (with a low-point anchor). For this analysis, the χ^2 was statistically significant (47.26; $P = 0.00$), and four presurgical variables (age, income, litigation, levels fused) had a Wald value large enough to achieve statistical significance. The corresponding risk ratios (Expβ) for each of the variables also are presented.

The results from another multiple-logistic regression predicting the Stauffer-Coventry final outcome classifications (good or fair *vs* poor) from the nine-variable presurgical model are presented in Table 6. The χ^2 for this model was statistically significant (22.48; $P = 0.00$). Age and litigation were statistically significant.

The results from a simultaneous-entry, multiple-linear regression of the nine-variable presurgical model on the Roland and Morris DQ is presented in Table 7. The model accounted for a significant amount of variance in DQ total score. The presurgical variables of age at the time of surgery, litigation, and prior low back operations were statistically significant predictors.

Tables 8 to 11 contain the regressions of the nine presurgical variables on the SF-20 subscales. The regression models for the SF-20 mental health and current health perceptions subscales are not presented because these equations were not statistically significant. The most consistent predictors across most SF-20 subscales were prior low back surgeries, income, age at time of surgery, and litigation. Depression was predictive of the SF-20 mental health and pain subscale. The number of vertebral levels fused was predictive of disability status. Importantly, time delay from injury to fusion, smoking status at the time of surgery, and diagnostic severity of presurgical imaging studies were not predictive of any outcomes assessed in this study.

In summary, the most robust predictors across a multidimensional array of outcomes were number of prior low back operations, household income at the time of injury, age at the time of surgery, and litigation. Presurgical depression was predictive of the SF-20 mental health and pain scales. The number of vertebral levels fused predicted only postsurgical disability status.

Table 6. Logistic Regression Predicting Stauffer-Coventry Good/Fair *vs* Poor Categories, With Presurgical Variables as Predictors

Variable	β	P	Exp (B)	95% CI
Age (5-year increments from age 25 years)	0.3385	0.0162	1.4028	(1.06–1.85)
Household income ($100 increments from $0.00)	−0.1403	0.2082	0.8691	(0.70–1.08)
Time delay (1-year increments from 0 years)	0.1212	0.4340	1.1289	(0.83–1.53)
Smoking	−0.0163	0.9675	0.9838	(0.45–2.16)
Depression	0.2422	0.7432	1.2740	(0.30–5.44)
Litigation	0.9957	0.0168	2.7067	(1.20–6.12)
Prior low back operations	0.5046	0.0979	1.6563	(0.91–3.01)
Levels fused	0.0284	0.9339	1.0288	(0.53–2.01)
Imaging severity rating (4-point increments from 0)	−0.2027	0.3749	0.8165	(0.52–1.28)
Constant	−2.8420	0.0600		

Exp (B) = estimated odds ratio; CI = critical interval for exp (B).

Table 7. Simultaneous-Entry Multiple Regression: Predicting the Roland and Morris (DQ) Total Score, With Presurgical Variables as Predictors

	Coefficients			
	Unstandardized Coefficients		Standardized Coefficients	
Variable	β	*SE*	β	*P*
Age	0.157	0.066	0.214	0.020
Household income	−0.005	0.003	−0.129	0.144
Time delay	−0.000	0.001	−0.020	0.820
Smoking	1.176	1.209	0.084	0.332
Depression	1.681	2.112	0.070	0.428
Litigation	2.496	1.250	0.175	0.048
Prior low back operations	2.028	0.864	0.210	0.021
Levels fused	0.297	1.020	0.025	0.771
Imaging severity rating	−0.204	0.141	−0.128	0.150
Constant	0.209	4.815		0.966

SE = standard error; R = .41; *P* = 0.008.

Table 9. Simultaneous-Entry Multiple Regression: Predicting the Short-Form 20 (SF-20) Role Functioning Subscale, With Presurgical Variables as Predictors

	Coefficients			
	Unstandardized Coefficients		Standardized Coefficients	
Variable	β	SE	β	*P*
Age	−0.948	0.407	−0.210	0.022
Household income	0.029	0.020	0.125	0.154
Time delay	−0.002	0.007	−0.022	0.800
Smoking	−4.383	7.430	−0.051	0.556
Depression	−15.939	12.986	−0.108	0.222
Litigation	−13.486	7.684	−0.153	0.082
Prior low back operations	−12.677	5.313	−0.212	0.019
Levels fused	−0.444	6.269	−0.006	0.944
Imaging severity rating	0.619	0.866	0.063	0.476
Constant	120.212	29.604		0.000

SE = standard error; R = .41; *P* = 0.006.

■ Discussion

The purpose of this study was to examine presurgical predictors and outcomes for posterolateral lumbar fusion in injured Utah workers. The authors also sought to determine whether the biopsychosocial model was a useful approach to predicting outcomes of lumbar fusion. Follow-up information was gathered for 130 patients (70%) an average of 5 years after surgery. Reported arthrodesis, reoperation, and disability rates for the follow-up sample were 74%, 24%, and 25%, respectively. A sizable minority of the patients (41%) thought that their overall quality of life had not improved or had worsened as a result of lumbar fusion, and 43% received Roland and Morris DQ scores in the "poor" range. Analysis of the SF-20 health subscales showed that patients who underwent fusion perceived their health status as substantially worse than that perceived by medical patients or general nonpatients.

Some of the outcome measures assessed in the current study can be compared with similar measures used in the study by Franklin et al.[5] A comparison of identical patient satisfaction items across the two studies showed that whereas a substantial proportion of both subject samples were dissatisfied with their surgical results, the Utah sample showed somewhat higher rates of satisfaction than the Washington sample. Overall, the two samples were quite comparable in terms of gender, average age at the time of surgery, typical preoperative diagnosis, and distribution of the vertebral levels fused. However, they differed in percentage of patients with prior low back surgeries (Washington = 61% *vs* Utah = 45%). Importantly, increased number of prior low back operations was shown to be a significant predictor of poor patient outcomes in both studies. Therefore, the better lumbar fusion outcomes in Utah might be partly because of fewer patients with prior back surgeries.

Table 8. Simultaneous-Entry Multiple Regression: Predicting the Short-Form 20 (SF-20) Physical Functioning Subscale, With Presurgical Variables as Predictors

	Coefficients			
	Unstandardized Coefficients		Standardized Coefficients	
Variable	β	SE	β	*P*
Age	−0.538	0.287	−0.172	0.064
Household income	0.018	0.014	0.114	0.204
Time delay	−0.001	0.005	−0.020	0.825
Smoking	−3.776	5.243	−0.063	0.473
Depression	−6.415	9.162	−0.063	0.485
Litigation	−9.367	5.421	−0.154	0.087
Prior low back operations	−8.258	3.749	−0.200	0.030
Levels fused	−1.385	4.423	−0.027	0.755
Imaging severity rating	0.404	0.611	0.059	0.510
Constant	89.754	20.887		0.000

SE = standard error; R = .37; *P* = 0.038.

Table 10. Simultaneous-Entry Multiple Regression: Predicting the Short-Form 20 (SF-20) Social Functioning Subscale, With Presurgical Variables as Predictors

	Coefficients			
	Unstandardized Coefficients		Standardized Coefficients	
Variable	β	SE	β	*P*
Age	−0.685	0.331	−0.191	0.041
Household income	0.033	0.016	0.183	0.043
Time delay	0.002	0.006	0.027	0.760
Smoking	−10.000	6.046	−0.146	0.101
Depression	−5.372	10.567	−0.046	0.612
Litigation	−3.342	6.252	−0.048	0.594
Prior low back operations	−8.310	4.323	−0.175	0.057
Levels fused	−1.253	5.101	−0.021	0.806
Imaging severity rating	−0.423	0.705	−0.054	0.549
Constant	118.445	24.088		0.000

SE = standard error; R = .36; *P* = 0.044.

Table 11. Simultaneous-Entry Multiple Regression: Predicting the Short-Form 20 (SF-20) Pain Subscale, With Presurgical Variables as Predictors

	Coefficients			
	Unstandardized Coefficients		Standardized Coefficients	
Variable	β	SE	β	P
Age	0.149	0.219	0.063	0.496
Household income	−0.021	0.011	−0.177	0.049
Time delay	−0.005	0.004	−0.115	0.198
Smoking	−3.578	3.996	−0.079	0.372
Depression	15.315	6.984	0.196	0.030
Litigation	4.115	4.132	0.089	0.321
Prior low back operations	5.564	2.857	0.177	0.054
Levels fused	0.935	3.371	0.024	0.782
Imaging severity rating	0.039	0.466	0.007	0.934
Constant	40.137	15.920		0.013

SE = standard error; R = .37; $P = 0.031$.

The solid fusion rates also varied, with the Utah sample having a lower fusion rate (71.9%) than the Washington sample (84.6%). This finding is difficult to interpret because the Utah sample had overall better outcomes than the Washington sample. However, the correlations between solid fusion and other patient outcomes in the current study were low (average r = 0.19). Other presurgical variables of interest in the current study (socioeconomic status, weekly income, presence of a lawyer, smoking, and diagnostic severity) were not assessed in the Franklin et al[5] study. It is plausible that the differences between the two samples in terms of these other presurgical variables, particularly the presence of a lawyer and weekly income, or other variables not included in either study, may further account for outcome discrepancies.

The current study demonstrated that a number of patient outcomes could be predicted from presurgical data. Variables consistently predictive of worse lumbar fusion outcomes included greater number of prior low back operations, low household income at the time of injury, older age at the time of surgery, and lawyer involvement at the time of surgery. Psychological depression also was a significant predictor of poor outcomes, but on a less consistent basis. Prior research has identified older age and increased number of prior low back operations as risk factors for poor lumbar fusion outcomes.[5,16,24] The deleterious effects of litigation on nonfusion back surgery outcomes have been established,[20,25,28] and the current study generalizes the predictive import of this variable to posterolateral lumbar fusion. Household income was shown to be predictive in only one prior nonfusion back surgery study,[6] and the present investigation is the first lumbar fusion study to demonstrate its predictive utility. Low back pain studies[8,17] and a nonfusion low back surgical study[20] have shown depression to be a consistent risk factor for poor outcomes. The present investigation is one of the first lumbar fusion studies to demonstrate a correlation between presurgical depression and worse patient outcomes.

Several hypotheses have been offered to explain why these constructs predict lumbar fusion outcomes. Older age, for example, is thought to be associated with more severe preoperative spinal pathology, and therefore worse functional lumbar fusion outcomes.[5,8,24] However, this explanation was not supported in the current study because diagnostic severity was not predictive of patient outcomes. An alternative explanation maintains that older persons might lack the biophysical resources to heal as efficiently as younger persons after lumbar spine fusion. Individuals near retirement age also might use their back surgery as a justification for seeking early retirement, thus becoming psychologically invested in adopting the disabled role.

In terms of prior back operations, it may be assumed that multiple prior operations cause increased scar tissue in the lumbar spine, which results in worse functional patient outcomes. It also has been suggested that individuals undergoing multiple operations are in poorer physical condition at surgery, thus impeding rehabilitation efforts and return to work.[17,24] A clinical finding from the current study shows that multiple prior spine operations often result in increased psychosocial stressors, which may detract from a patient's rehabilitation efforts and exacerbate the risk for poor outcomes.

Lower household income may be a risk factor for poor outcome simply because it is an indicator of limited resources available to the patient both before surgery and during rehabilitation.[6] Patients with lower incomes might delay seeking medical care for a back injury, might not be able to afford optimal rehabilitation, might return to work prematurely, or might return to the physically demanding job that initially resulted in their back injury. Alternatively, people earning a low wage at the time of injury might have a financial incentive not to return to work because they may be provided with a similar wage when on disability.

Patients who hire lawyers may be more invested in "proving" disability with a secondary motive of winning a large legal settlement.[6,8] Alternatively, these patients may have more severe spinal pathologies before surgery, although this was not the case in the current study. Psychological depression is related to increased stress, and both factors have been linked to decreased immune system function and consequently more prolonged healing after surgery.[26] Individuals experiencing depression also are less likely to call on social supports or follow a plan of rehabilitation.[26]

In general, the biopsychosocial model was supported by the predictive analyses presented in this article. It appears that presurgical biologic factors (age, prior low back operations) and psychosocial factors (lawyer involvement, low household income, depression) are predictive of outcomes from posterolateral lumbar fusion. It appears that poor lumbar fusion outcomes could be potentially minimized if selection techniques were used to identify the patients at highest risk. The current study identified five such presurgical risk factors that could be

integrated into a selection and intervention program. For example, high-risk patients can be educated about their likelihood for recovery, and in some cases may choose not to undergo the surgery. Alternatively, patients experiencing depression may wait until their emotional condition improves before undertaking the stresses of surgery. High-risk patients who proceed with surgery might participate in a more rigorous and closely monitored rehabilitation. These intervention strategies deserve careful attention in future investigations.

The limitations of this study include its retrospective and correlational nature. There is a clear need to validate the current findings *via* randomized controlled trials in which nonexperimental variation can be controlled.

■ Key Points

- A retrospective cohort study examining outcomes of posterolateral lumbar fusion in compensated workers (n = 185) is presented.
- A medical chart review was conducted, and outcomes were assessed including solid fusion, reoperation, disability rates, satisfaction, functional status, and general health status.
- Follow-up surveys were completed with 130 patients (70%), and outcomes were inconsistent across measures.
- Presurgical predictors of outcomes were prior low back operations, income at the time of injury, age, litigation, and depression.
- It is concluded that screening for presurgical sociodemographic variables may be important for surgical decisions and rehabilitation of compensated lumbar fusion patients.

References

1. Bernard TN. Repeat lumbar spine surgery: Factors influencing outcome. Spine 1993;18:2196–2200.
2. Deyo RA, Cherkin D, Conrad D, et al. Cost, controversy, crisis: Low back pain and the health of the public. Annu Rev Public Health 1991;12:141–56.
3. Doxey NV, Dzioba RB, Mitson GL, et al. Predictors of outcome in back surgery candidates. J Clin Psychol 1988;44:611–22.
4. Dzioba RB, Doxey NC. Predictors of outcome of repeat back surgery following industrial injuries. Proc 51st Annu Meet Am Acad Orthop Surg 1984; 407:155–61.
5. Franklin GM, Haug J, Heyer NJ, et al. Outcome of lumbar fusion in Washington State workers' compensation. Spine 1994;17:1897–903.
6. Frymoyer JW. Predicting disability from low back pain. Clin Orthop 1992; 279:101–9.
7. Frymoyer JW, Cats-Baril WL. An overview of the incidences and costs of low back pain. Orthop Clin North Am 1991;22:263–71.
8. Frymoyer JW, Cats-Baril WL. Predictors of low back pain disability. Clin Orthop 1987;221:89–98.
9. Frymoyer JW, Pope MH, Clemens JH, et al. Risk factors in low back pain. J Bone Joint Surg [Am] 1983;65:213–18.
10. Hanley EN. The indication for lumbar spinal fusion with and without instrumentation. Spine 1995;20:143s–53s.
11. Katz JN, Spratt KF, Andersson GBJ, et al. Epidemiology introduction: 1995 focus issue meeting on fusion. Spine 1995;20:76s–7s.
12. Mayer T, McMahon J, Gatchel RJ, et al. Socioeconomic outcomes of combined spine surgery and functional restoration in workers' compensation spinal disorders with matched controls. Spine 1998;23:598–606.
13. Oostdam EM, Duivenvoorden HJ. Predictablity of the result of surgical intervention in patients with low back pain. J Psychosom Res 1983;27:273–81.
14. Oostdam EM, Duivenvoorden HJ, Pondaag W. Predictive value of some psychological tests on the outcome of surgical intervention in low back pain patients. J Psychosom Res 1981;25:227–35.
15. Owings MF, Kozak LJ. Ambulatory and inpatient procedures in the United States, 1996. National Center for Health Statistics. Vital Health Stat 13(139), 1998.
16. Pfeiffer M, Griss P, Haake M, et al. Standardized evaluation of long-term results after anterior interbody fusion. Eur Spine J 1996;5:299–307.
17. Polatin PB, Gatchel RJ, Barnes D, et al. A psychosocomedical prediction model of response to treatment by chronically disabled workers with low back pain. Spine 1988;14:956–61.
18. Roland M, Morris R. A study of the natural history of low back pain: Part I. Development of a reliable and sensitive measure of disability in low back pain. Spine 1983;8:141–4.
19. Roland M, Morris R. A study of the natural history of low back pain: Part II. Development of guidelines for trials of treatment in primary care. Spine 1983;8:145–50.
20. Sorenson LV, Mors O, Skovlund O. A prospective study of the importance of psychological and social factors for the outcome after surgery in patients with slipped lumbar disk operated upon for the first time. Acta Neurochir 1987;88:119–25.
21. Stauffer R, Coventry MB. Anterior interbody lumbar spine fusion. Analysis of Mayo Clinic series. J Bone Joint Surg [Am] 1972;54:756–68.
22. Stewart AL, Hays RD, Ware JE. The MOS short-form general health survey: Reliability and validity in a patient population. Med Care 1988;26:724–35.
23. Stewart AL, Ware JE, eds. Measuring Functioning and Well-Being: The Medical Outcomes Study Approach 1992. Durham, London: Duke University Press.
24. Stewart G, Sachs BL. Patient outcomes after reoperation on the lumbar spine. J Bone Joint Surg [Am] 1996;78:706–11.
25. Taylor ME. Return to work following back surgery: A review. Am J Ind Med 1989;16:79–88.
26. Taylor SE. Health Psychology. 4th ed. New York: McGraw-Hill, 1999.
27. Turner JA, Ersek M, Herron L, et al. Patient outcomes after lumbar spinal fusions. JAMA 1992;268:907–11.
28. Uomoto JM, Turner JA, Herron LD. Use of the MMPI and MCMI in predicting outcome of lumbar laminectomy. J Clin Psychol 1988;44:191–7.
29. Waddell G, McCulloch, JA, Kummel E, et al. Nonorganic physical signs in low-back pain. Spine 1979;5:117–25.
30. Watkins RG, O'Brien JP, Draugelis R, et al. Comparison of preoperative and postoperative MMPI data in chronic back patients. Spine 1986;11:385–90.
31. Wifling FJ, Klonoff H, Kokan P. Psychological, demographic, and orthopaedic factors associated with prediction of outcome of spinal fusion. Clin Orthop 1973;90:153–60.

Address reprint requests to

M. Scott DeBerard, PhD
Department of Psychology
Utah State University
2810 Old Main Hill
Logan, UT 84322-2810
E-mail: sdeberard@coe.usu.edu

■ Appendix

Imaging Study Diagnostic Severity Rating Form

Patient's I.D. Number_____ Latest Preoperative Films: Plain___CT ___ MRI___

L2-3 LEVEL				
Degenerative Changes Disc-Facet	None	Mild	Moderate (Dessication-Narrowing-Facet Changes)	Severe (Vacuum-Modic-Facet Overgrowth)
Disc Bulges	None	I Bulging-No Abutment	II (Abutting-Crowding of Nerves)	III (Displacing Nerve Tissue)
Listhesis-Anterior or Posterior	None or 2mm or less		<5mm	5mm or more
Lysis	None	Present		
Stenosis-(Formina or Far Lateral)	None	Mild	Moderate	Severe
Stenosis (Central or Spinal)	None	Mild	Moderate	Severe
Discography	Normal	Discordant (Atypical Pain or Typical Pain with Normal Anatomy)	Concordant (Typical Pain with Abnormal Anatomy)	

L3-4 LEVEL				
Degenerative Changes Disc-Facet	None	Mild	Moderate (Dessication-Narrowing-Facet Changes)	Severe (Vacuum-Modic-Facet Overgrowth)
Disc Bulges	None	I Bulging-No Abutment	II (Abutting-Crowding of Nerves)	III (Displacing Nerve Tissue)
Listhesis-Anterior or Posterior	None or 2mm or less		<5mm	5mm or more
Lysis	None	Present		
Stenosis-(Formina or Far Lateral)	None	Mild	Moderate	Severe
Stenosis (Central or Spinal)	None	Mild	Moderate	Severe
Discography	Normal	Discordant (Atypical Pain or Typical Pain with Normal Anatomy)	Concordant (Typical Pain with Abnormal Anatomy)	

L4-5 LEVEL				
Degenerative Changes Disc-Facet	None	Mild	Moderate (Dessication-Narrowing-Facet Changes)	Severe (Vacuum-Modic-Facet Overgrowth)
Disc Bulges	None	I Bulging-No Abutment	II (Abutting-Crowding of Nerves)	III (Displacing Nerve Tissue)
Listhesis-Anterior or Posterior	None or 2mm or less		<5mm	5mm or more
Lysis	None	Present		
Stenosis-(Formina or Far Lateral)	None	Mild	Moderate	Severe
Stenosis (Central or Spinal)	None	Mild	Moderate	Severe
Discography	Normal	Discordant (Atypical Pain or Typical Pain with Normal Anatomy)	Concordant (Typical Pain with Abnormal Anatomy)	

L5-S1 LEVEL				
Degenerative Changes Disc-Facet	None	Mild	Moderate (Dessication-Narrowing-Facet Changes)	Severe (Vacuum-Modic-Facet Overgrowth)
Disc Bulges	None	I Bulging-No Abutment	II (Abutting-Crowding of Nerves)	III (Displacing Nerve Tissue)
Listhesis-Anterior or Posterior	None or 2mm or less		<5mm	5mm or more
Lysis	None	Present		
Stenosis-(Formina or Far Lateral)	None	Mild	Moderate	Severe
Stenosis (Central or Spinal)	None	Mild	Moderate	Severe
Discography	Normal	Discordant (Atypical Pain or Typical Pain with Normal Anatomy)	Concordant (Typical Pain with Abnormal Anatomy	

Point of View

John Frymoyer, MD
Professor Emeritus
University of Vermont
Burlington, Vermont

Some readers will remember the refrain of a Vietnam-era song, which inquired, "When will [we] ever learn?" Over the past 40 years, a massive literature has accumulated that emphasizes the pschycosocial characteristics of chronic disability and the poor outcomes in the chronically disabled from a variety of treatments, including surgical and nonsurgical interventions. It also has been learned that surgical outcomes are far more predictable when the pathology for the patient's symptoms are well established by accurate preoperative diagnosis including clinical history, physical examination, and confirming imaging studies. Despite the development of new fusion techniques, the initial enthusiasm for the technique inevitably has been followed by the reality that the new mousetrap is not the solution to the problem of disability. These collective observations have been incorporated into predictive risk models for disability and surgical outcomes that have been shown to have a high degree of sensitivity and specificity.

In the current article, "Outcomes of Posterolateral Lumbar Fusion in Utah Patients Receiving Workers' Compensation," the investigators have used a biopsychosocial model that incorporates the relevant clinical, and psychological indicators. Again, this population-based study demonstrates the poor outcomes that ensue from spinal fusion among compensated workers. Their results confirm prior studies on the outcomes of lumbar fusion in worker's compensation patients in Washington State.

The critics will question the retrospective design and marginal retrieval rate (70%). Others may ask, "Could the results be improved by better fusion techniques such as combined anteroposterior fusion?" A more useful question is this: "How can we better predict the patients who will have a good result?" The authors have given us further insights that should permit the development of better predictive risk models, which can be used by individual surgeons trying to give the best care to this complex patient cohort. But the basic question raised by this study is inescapable: "When will [we] ever learn?"