PSYCHOLOGY, HEALTH & MEDICINE, VOL. 7, NO. 4, 2002 BrunnerRoutledge Taylor&Francis healthsciences

Pre-surgical psychological screenings for lumbar fusion: a look at real world practice

M. S. DEBERARD,[1] K. S. MASTERS,[1] A. L. COLLEDGE,[2] R. L. SCHLEUSENER[3] & J. D. SCHLEGEL[3]

[1]*Utah State University,* [2]*Labor Commission of Utah &* [3]*University of Utah Medical School, Salt Lake City, USA*

Abstract *Candidates for lumbar fusion are often referred for psychological evaluation prior to surgery in order to identify psychosocial issues that may complicate outcomes and provide recommendations for how to ameliorate these potential problems. The long-term goal of such evaluations is to optimize successful surgical outcomes. The purpose of this study was to examine a cohort of compensated lumbar fusion patients from Utah who underwent pre-surgical psychological evaluation (n = 29) and to compare their surgical outcomes with patients who did not have these evaluations (n = 115). A retrospective cohort design was utilized that included a pre-surgical medical chart review and a post-surgical telephone survey with patients at least two years following surgery. Results indicated that patients referred for psychological evaluations had higher levels of pre-surgical alcohol use and depression, and following surgery had longer recovery times, higher medical and compensation costs and disability rates than other fusion patients. The two cohorts did not differ in terms of patient satisfaction, back pain-related impairment or general health status. The significance of these findings for surgical decisions, rehabilitation interventions and professional psychological practice is discussed.*

Introduction

The rates of lumbar fusion surgery among injured workers are increasing in the USA (Taylor *et al.*, 1994). This is an expensive procedure that requires a lengthy rehabilitation period, often up to at least six months for a solid fusion to form and longer for maximum functional recovery. In Utah, the average medical and compensation costs for injured workers undergoing lumbar fusion are $59,000 and the average time for maximal medical improvement is 364 days (DeBerard, 1998). Total expenditures for 203 compensated lumbar fusion patients from Utah who had surgeries between 1990 and 1995, including long-term disability settlements, were $12 million (DeBerard, 1998). Despite these enormous costs, 23% of posterolateral lumbar fusion patients in Utah never return to work (DeBerard *et al.*, 2001) and disability rates for compensated lumbar fusion in other states are even higher (e.g. Franklin *et al.*, 1994).

Given the high costs and risk for chronic disability associated with this procedure, investigators have sought to identify pre-surgical medical and psychosocial factors that

Address for correspondence: M. Scott DeBerard, Department of Psychology, Utah State University, 2810 Old Main Hill Logan, Utah 84322–2810, USA. Tel: + 1 435 797 1462; Fax: + 1 435 797 1448; E-mail: sdeberard@coe.usu.edu

ISSN 1354-8506 print/ISSN 1465-3966 online/02/040411-14 © Taylor & Francis Ltd
DOI: 10.1080/1354850021000015230

predict surgical outcomes (e.g. Block & Callewart, 1999; DeBerard *et al.*, 2001; Franklin *et al.*, 1994; Turner *et al.*, 1992). The search for medical predictors has been perplexing as studies have found that pre-surgical diagnoses are not predictive of fusion outcomes (DeBerard *et al.*, 2001; Franklin *et al.*, 1994; Turner *et al.*, 1992). For example, DeBerard and colleagues (2001) found that a diagnostic severity index based upon pre-surgical imaging studies (MRI, CT) did not predict fusion outcomes. Other pre-surgical demographic and psychosocial factors have been shown to be more consistently predictive of patient outcomes. For example, older age, litigation, increased number of prior low back surgeries, low income, compensation, increased time of work disability, and depression have all been shown to be predictive of lumbar fusion outcomes (DeBerard *et al.*, 2001; Franklin *et al.*, 1994; Taylor, 1989; Turner *et al.*, 1992; Wifling *et al.*, 1973). Screening patients for such factors should allow for identification of high-risk patients, appropriate pre-surgical interventions and post-operative follow-up.

One means for identifying these important contraindications prior to surgery is to refer patients to a clinical psychologist for a pre-surgical psychological evaluation (Block & Callewart, 1999). The criteria for referring patients for psychological evaluation are, however, not standardized. In Utah, surgeons typically make a referral based upon suspected drug/alcohol use, depression, anxiety or malingering. The actual referral rate for psychological evaluations is unknown, and there is little consensus regarding the protocol for these evaluations. Most often, psychological evaluations seek to uncover potential complicating factors such as those mentioned above (Block & Callewart, 1999). Typically, information from the report is reviewed by the surgeon and corollary providers and is used in deciding whether to proceed with the operation or implement an intervention to increase the chances for surgical success. It is presumed that if a patient has surgery following a psychological evaluation that psychosocial factors that may complicate recovery were ruled out or that such issues were adequately treated and resolved prior to surgery. Thus, we would expect if the pre-surgical psychological evaluation process is conducted appropriately, those who receive psychological evaluation would be at no greater risk for poor surgical outcome than other patients who were not evaluated. However, this hypothesis has yet to be empirically addressed.

The purpose of this study was to compare surgical outcomes for a cohort of compensated lumbar fusion patients who received pre-surgical psychological evaluations versus a matched group who did not receive a psychological evaluation. The following research questions were assessed: (1) what are the referral rates for psychological evaluation among compensated lumbar fusion patients?; (2) what are the pre-surgical biopsychosocial characteristics of lumbar fusion patients referred for psychological evaluations and are these characteristics comparable to a group of lumbar fusion patients who did not receive psychological evaluations?; and (3) what are the biopsychosocial outcomes for lumbar fusion patients who received psychological evaluations and are these outcomes consistent with fusion patients who did not receive pre-surgical psychological evaluations?

Method

Study design

This was a retrospective cohort study that used pre-surgical information documented in patient medical records and information from a telephone outcome survey conducted with patients at least two years following their surgeries. The Institutional Review Boards from the University of Utah Medical School and Utah State University approved this study.

Participants

All patients who had undergone posterolateral lumbar fusion surgery from 1990 to 1995, who did not have a pre-surgical diagnosis of vertebral fracture, and who were at least two years post-surgery at the time of follow-up were eligible for inclusion ($N = 203$). Patients were identified via Physician's Current Procedural Terminology (CPT) codes as established by the American Medical Association and located in Workers' Compensation Fund of Utah (WCFU) databases. For each patient, lumbar fusion surgery was performed because of a verified workplace low-back injury. Following primary data collection, CPT codes were also used to identify patients who completed a pre-surgical psychological evaluation ($n = 29$). Thus, data abstractors were blinded to psychological evaluation group membership during the primary data collection process and those conducting telephone outcome follow-up assessment were also unaware of group membership status.

Pre-surgical information

Pre-surgical medical record data were gathered via independent and standardized review of medical chart information and review of WCFU computer databases. No one involved in the research had any clinical contact with the study participants. Data abstractors were graduate students trained to complete a standard coding instrument. Where possible, the actual psychological reports were obtained from the medical file. The biopsychosocial variables coded included: age at time of surgery; sex; smoking status at time of surgery (coded yes/no); educational status at time of injury; surgeon's diagnosis of current depressive disorder (coded yes/no; including major depressive disorder, dysthymia or adjustment disorder with depressed mood); surgeon's assessment of alcohol use (coded yes/no); household income at time of injury; litigation as defined by patient private lawyer involvement in the compensation case at the time of surgery (coded yes/no); number of prior workplace injuries; and number of prior low back operations. Medical variables coded included number of vertebral levels to be fused, instrumentation, pre-surgical diagnostic severity score based upon independent review of pre-surgical radiology reports by three of the authors (ALC, RLS, JDS) and pre-operative diagnoses. The pre-surgical diagnostic severity measure allows for coding of six major types of common lumbar spinal diagnoses (degenerative changes, disc bulge, listhesis, lysis, foraminal stenosis and central stenosis) and provides a graduated scale for each diagnoses from none (0) to severe (3). This scale is applied to four vertebral levels (L2–3, L3–4, L4–5, L5-S1) and a total score is achieved by adding the single scores for each of the assessed lumbar levels. A high degree of correspondence was achieved by the three raters (92% agreement). Discrepancies (8%) were resolved by conference. The measure was published in a prior article (DeBerard *et al.*, 2001).

Medical and compensation outcomes

The medical outcome variables were arthrodesis (solid fusion) rates, re-operation rates and time to maximal medical improvement (MMI). Arthrodesis was assessed by reviewing the operating surgeon's post-operative findings in the patient medical chart. For all patients, operating surgeons routinely reviewed radiographs and documented the progression of bone consolidation. The operating surgeon's final determination of arthrodesis was coded as a dichotomous variable (yes/no). Radiologists' reports of patient radiographs were found in the

medical chart and were independently reviewed by two orthopaedic surgeons (RLS and JDS) and an occupational physician (ALC). Each reviewer was blinded to psychological referral status. The operating surgeons' final opinion of arthrodesis was typically determined and documented in the chart at the time of patient MMI. Re-operation was coded if the patient received another low back operation for the original injury.

MMI is defined in Utah as the point in time at which an injured worker has realized the maximum amount of improvement from an injury given the current standard of medical care. The operating surgeon is responsible for monitoring the patient's status and declaring the date of MMI. At the point of MMI, the patients' functional capabilities and degree of impairment are assessed and the surgeon can recommend return to previous work, return to modified or lighter work, job retraining, or no return to work. Compensation outcomes consisted of the total medical and compensation costs from injury to MMI, including long-term disability settlements. Disability status at the time of follow-up was determined if patients were currently receiving total disability benefits for their back condition. Medical and compensation outcome data were gathered from WCFU computer databases and medical records.

Patient telephone outcome survey

An objective outcome instrument, including a script for telephone interviewers, was created. This instrument included four patient satisfaction items (quality of life improvement as a result of lumbar fusion surgery, perceived back or leg pain improvement, satisfaction with back condition at time of follow-up, would patient choose to have surgery again) drawn from a recent study of compensated lumbar fusion patients (Franklin *et al.*, 1994).

The Roland & Morris Back Pain Disability Questionnaire (DQ; Roland & Morris, 1983a; 1983b), a 24-item self-report measure designed to evaluate dysfunction associated with low back pain, was also used. The Stauffer-Coventry Index (Stauffer & Coventry, 1972) was selected as a clinical surgical outcome measure. This index, widely used for assessing low back surgical outcomes (e.g. Boos *et al.*, 1991; 1992; Oostdam & Duivenvoorden, 1983; Oostdam *et al.*, 1981; Uomoto *et al.*, 1988), was used as an outcome measure in two relatively recent reviews (Taylor, 1989; Turner *et al.*, 1992). Designed for after surgery administration, it consists of four multiple response 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: (a) good: 76–100% relief in leg and back pain, return to previous work status, minimal or no restriction of physical activities, occasional mild analgesics or no analgesics; (b) fair, 26–75% relief of leg and back pain, return to lighter work, moderate restrictions of physical activities, regular use of non-narcotic analgesics; and (c) poor: 0–25% relief of leg and back pain, no return to work following surgery, severe restrictions of physical activities, occasional or regular use of narcotic analgesics.

The Short Form Health Survey-20 (SF-20; Stewart & Ware, 1992; Stewart *et al.*, 1988), a 20-item general health survey, was used to assess six general dimensions of health: (1) physical functioning (the extent to which health interferes with a variety of activities); (2) role functioning (the extent to which health interferes with usual daily activities); (3) social functioning (the extent to which health interferes with normal social activities); (4) mental health (general mood); (5) current health perceptions (overall ratings of current general health); and (6) pain (the extent of bodily pain during the past month).

Procedures for telephone outcome survey

The initial contact occurred via a letter sent to patients' most recent known address. This letter, sent to all patients who met inclusion criteria, explained the purposes of the study, procedures and assurances of confidentiality, with a request for their participation. A self-addressed, stamped card was included. Patients who mailed cards back were immediately contacted for the telephone survey. Patients who did not send back cards were sent as many as two additional contact letters requesting their participation. If patients still did not respond, a telephone contact was attempted, and if the patient verbally consented, the telephone survey was completed at that time.

Data analysis

Data were analyzed using SPSS for Windows (Version 11.0). Univariate statistical tests were two-sided. P values < 0.5 were considered statistically significant. Statistical analyses consisted of: (1) F tests to compare the differences between group means; (2) chi-square tests of independence for assessing differences in categorical variables between groups, and (3) Mann-Whitney Rank Sum Tests for determining differences between groups when ordinal data were involved. Measures of effect size (e.g. standardized mean difference/Phi-coefficient) were also calculated.

Results

Patient and follow-up data

A total of 203 patients (82.3% men) were included in the study cohort. Patients underwent surgery between 23 August 1990, and 21 April 1995. Ninety-six per cent were Caucasian, 3% were Hispanic, 0.5% were Native American and 0.5% were Asian-American. Pre-operative diagnoses were as follows: disc herniation (70.4%), degenerative disc disease (46.3%), segmental instability (28.6%), spondylolisthesis (17.8%), spinal stenosis (17.7%) and pseudoarthrosis (2.5%). Subjects could receive multiple pre-surgical diagnoses. All received posterior or posterolateral lumbar fusions. Instrumentation was used in 83.7% of cases, with pedicle screw fixation used in 93% of instrumented cases. A total of 28 surgeons were involved in the 203 operations. The overall follow-up rate for the patient outcome survey was 70.9% ($n = 144$). Average time to follow-up was 4.6 years (range = 2.1 to 7.3 years). A series of univariate and chi-square analyses were conducted to determine whether follow-up respondents versus non-respondents differed on the 13 pre-surgical characteristics. These analyses were not statistically significant and it was concluded that results of the outcome survey were not differentially biased between respondents and non-respondents due to measured pre-surgical characteristics.

Referral rates for psychological evaluation

A total of 41 out of 203 patients (20.2%) received psychological evaluations prior to surgery. For those participating in the follow-up, 29 out of 144 patients (20.1%) received psychological evaluations. There were no differences in rates of psychological referral between surgeons, $\chi^2(22, N = 144) = 13.48$, $p = 0.92$.

Pre-surgical characteristics of cohorts

A series of univariate and chi-square tests were conducted that compared 13 pre-surgical characteristics between those referred for a psychological evaluation and those not referred. Table 1 contains the descriptive statistics, results of the univariate F tests, chi-square tests and effect size differences for the pre-surgical characteristics of psychological evaluation versus non-psychological evaluation patients. The only statistically significant differences between the two groups on pre-surgical characteristics were depression and alcohol use at time of surgery. The psychological evaluation cohort had a higher prevalence of depression and alcohol use than did the non-psychological evaluation group. A chi-square analysis comparing the two cohorts on pre-surgical diagnoses was conducted and was not statistically significant.

Medical and compensation outcomes

The arthrodesis rate was 73.4% ($n = 149$) for the total study sample ($N = 203$) and 76.4% ($n = 110$) for the follow-up sample ($n = 144$). Arthrodesis rates did not vary between respondents and non-respondents. The psychological report cohort arthrodesis rate was 83.0% ($n = 24$) versus 74.8% ($n = 86$) for the other patients (χ^2 (4, $N = 144$) = 0.82, $p = 0.36$). The re-operation rate for the total follow-up sample was 27.8% ($n = 40$). The psychological report cohort re-operation rate was 41.4% ($n = 12$) versus 24.3% ($n = 28$) for the other patients; this difference was not statistically significant (χ^2 (4, $N = 144$) = 3.35, $p = 0.07$). The mean average days to MMI for the total sample ($N = 203$) was 365 days (SD = 227.9), and for the follow-up sample ($n = 144$) it was 359.4 days (SD = 227.7). However, mean average days to MMI for the psychological report group were significantly greater (440 days, SD = 327.5) than for other patients (338 days, SD = 191.3); (F (1, 142), = 4.718, $p = 0.03$).

Twenty-four per cent (34) of the follow-up cohort were on permanent total disability. The permanent total disability rate for psychological evaluation patients was 44.8% (13) versus 18.3% (21) for other patients; this difference was statistically significant (χ^2 (4, $N = 144$) = 9.06, $p = 0.00$).

Average medical and compensation costs for the total sample ($N = 203$) were \$30,103 (SD = 16,398) and \$27,218 (SD = \$23,567), respectively. Average medical and compensation costs for the total follow-up sample ($n = 144$) were \$29,082 (SD = 14,940) and \$26,722 (SD = \$14,637), respectively. Average medical expenses were significantly higher for psychological evaluation patients (\$38,173, SD = \$20,377) versus other patients (\$26,789, SD = \$12,315); (F (1,142), = 14.74, $p = 0.00$). Average compensation expenses for psychological evaluation patients were also significantly higher (\$31,855, SD = \$16,299) versus other patients (\$25,428, SD = \$13,967); (F (1,142), = 4.58, $p = 0.03$). It is important to note that analyses of outcome variables by surgeon produced non-significant results on all measures.[1]

Patient outcome survey

Patient satisfaction items. Patient responses on the four patient satisfaction items did not differ across the psychological evaluation and non-psychological evaluation groups.

Roland and Morris DQ. The mean DQ score for the follow-up sample was 11.4 (SD = 6.96), with a minimum score of 0 and a maximum score of 24. The average score for the

Table 1. *Comparison of pre-surgical characteristics for patients with and without pre-surgical psychological evaluations*

Pre-surgical variable	Means or proportion: Psychological evaluation ($n = 29$)	Means or proportion: No psychological evaluation ($n = 115$)	Test value	Significance level	ES or ϕ
Age, years	38.12	38.71	$F = 0.09$	$p = 0.76$	− 0.06
Gender, %					
Male	79.3	85.2	$\chi^2 = 0.60$	$p = 0.44$	0.16
Female	20.7	14.8			
Reported as smoker, %	58.6	42.6	$\chi^2 = 2.39$	$p = 0.12$	0.13
Educational status, %					
< H.S	0.0	0.0	$\chi^2 = 3.98$	$p = 0.55$	0.17
Some HS	24.1	13.0			
HS/GED	48.3	41.7			
Technical	10.3	16.5			
Some college	13.8	20.9			
College graduate	3.4	7.0			
Graduate study	0.0	0.7			
Pre-surgical diagnosis of depression, %*	20.7	6.1	$\chi^2 = 6.01$	$p = 0.01$	0.20
Pre-surgical alcohol use, %*	62.1	31.3	$\chi^2 = 9.35$	$p = 0.00$	0.26
Household income at time of injury, $	417.02	424.03	$F = 0.032$	$p = 0.86$	− 0.04
Lawyer involvement in compensation case, %	51.7	33.8	$\chi^2 = 3.14$	$p = 0.08$	0.15
Prior work place injuries, %	65.5	48.7	$\chi^2 = 2.63$	$p = 0.11$	0.14
Prior low back operations, %					
None	62.1	51.3	$\chi^2 = 2.02$	$p = 0.57$	0.12
One	27.6	39.1			
Two	10.3	7.8			
3 or >	0.0	1.7			
Number of levels fused, %					
1 level	41.4	53.0	$\chi^2 = 1.48$	$p = 0.48$	0.10
2 level	55.2	42.6			
3 or More Levels	3.4	4.3			
Instrumentation used, %	89.7	85.2	$\chi^2 = 0.381$	$p = 0.54$	0.05
Mean pre-surgical imaging severity (higher value = greater severity)	7.00	6.73	$F = 0.090$	$p = 0.70$	0.07

*$p < 0.05$.
** ES or ϕ = effect size (SMD) or phi-coefficient.

psychological evaluation cohort was 12.34 (SD = 7.31) versus 10.87 (SD = 7.04) for the other patients; this difference was not statistically significant (F (1, 142), = 1.00, $p = 0.32$).

Stauffer-Coventry data. The four sub-scale values and aggregate ratings for the Stauffer-Coventry Index are presented in Table 2. The only statistically significant difference between the cohorts was for the employment status variable. Although the other sub-scale values and the aggregate score did not significantly differ, they all portray the tendency toward better function in the group that was not referred for a psychological evaluation, and these data produce moderate effect sizes.

Table 2. *Stauffer-Coventry Index: subscale scores and aggregate ratings*[1]

Pain relief ($U = 1391.49$, $z = -1.47$, $p = 0.14$)			Employment status ($U = 1182.50$, $z = -2.63$, $p = 0.01$)			Physical limitations ($U = 5221.02$, $z = -.64$, $p = 0.53$)			Medication usage ($U = 1475.48$, $z = -1.03$, $p = 0.30$)			Overall index rating* ($U = 1480.95$, $z = -1.04$, $p = 0.30$)		
Category	Psych. eval.	No psych. eval	Category	Psych. eval.	No psych. eval.	Category	Psych. eval.	No psych. eval.	Category	Psych. eval.	No psych. eval.	Category	Psych. eval.	No psych. eval.
Good (76–100% Improvement)	24.1% (7)	35.7% (41)	Good (return to previous work status)	10.3% (3)	31.3% (36)	Good (minimal or no restrictions)	13.8% (4)	19.1% (22)	Good (occasional or no use of mild analgesics)	37.9% (11)	47.8% (55)	Good	6.9% (2)	7.8% (9)
Fair (26–75% improvement)	41.4% (12)	41.7% (48)	Fair (return to lighter work)	51.7% (15)	48.7% (56)	Fair (moderate restrictions)	51.7% (15)	50.4% (58)	Fair (regular use on non-narcotic analgesics)	31.0% (9)	28.7% (33)	Fair	34.5% (10)	45.2% (52)
Poor (0–25% improvement)	34.5% (10)	22.6% (26)	Poor (no return to work)	37.9% (11)	20.0% (23)	Poor (severe restrictions)	34.5% (10)	30.4% (35)	Poor (occasional or regular use of narcotic analgesics)	31.0% (9)	23.5% (27)	Poor	58.6% (17)	47.0% (54)

[1] For each sub-scale, a Mann-Whitney statistic for rank ordinal differences across groups is provided under each sub-scale heading.
* Final classification based upon lowest rated single category.

SF-20. Table 3 contains the follow-up means and standard deviations for the SF-20. Table 3 also contains the normative means and standard deviations for the six SF-20 sub-scales for medical outpatients provided by Stewart *et al.* (1988). To further characterize differences between the follow-up cohorts and the normative sample, an effect size was calculated for each. Overall, both study cohorts demonstrated lower perceived health status compared to normative medical patients. This difference was evidenced by large (>0.8) to small (~ 0.2) effect sizes across a number of SF-20 health sub-scales. A MANOVA comparing psychological evaluation versus non-psychological evaluation patients across the SF-20 sub-scales was not statistically significant. The last four columns of Table 3 contain the percentage of patients falling in the poor health range for each of the sub-scales compared to medical outpatient and general population normative data (Stewart *et al.*, 1988). As can be seen, patients receiving lumbar fusion surgery, as a group, tend to demonstrate poorer health when compared to a general population and to other medical patient groups.

Discussion

Pre-surgical and outcome variables of compensated lumbar fusion patients with and without a pre-surgical psychological evaluation were examined. Approximately 20% of pre-surgical patients were referred for psychological evaluation. Those receiving pre-surgical psychological evaluation had higher frequencies of depression and alcohol use prior to surgery than other patients, but did not differ on several additional medical and demographic variables. Compared to those not receiving pre-surgical psychological evaluation, patients referred for evaluation evidenced statistically significant and meaningfully worse surgical and clinical outcomes on several important criteria. For example, psychological evaluation patients versus other patients showed a greater probability of post-surgical permanent total disability (44.8% versus 18.3%), lower return to work rates (10.3% versus 31.3%), longer surgery recovery time (time to MMI 440 days versus 338 days, or about 3.5 months longer) and higher medical ($38,173 versus $26,789) and compensation ($31,855 versus $25,428) costs. Psychological evaluation patients also demonstrated trends toward higher rates of re-operation (41.4% versus 24.3%) and reported less pain relief, increased physical limitations and increased medication usage than non-psychological evaluation patients. Arthrodesis did not differ significantly between the groups (83% psychological evaluation group versus 74.8%). General health status as measured by the SF-20 also did not differ as a function of pre-surgical psychological evaluation.

It is important to ascertain the reasons for the relatively poorer outcomes among those receiving pre-surgical psychological evaluations. Theoretically, in a best practice scenario, if a patient is referred for psychological evaluation it is presumed that any psychological factors that could jeopardize a successful outcome will be identified and remedied. These remedies could include treatments for disorders such as depression and anxiety, education regarding expectations for surgical outcome, directed vocational counselling, or substance abuse intervention to name a few. Alternatively, the psychologist may conclude that the patient presents with such a level of psychological risk that surgery is unlikely to be successful and should therefore be delayed or cancelled. In any event, if the evaluations are serving their intended purposes, those receiving psychological evaluation and actually undergoing surgery should not have worse outcomes than those who do not receive such evaluation. Unfortunately, the present study did not find this to be the case.

One possible explanation is that the surgeons ignored the psychologists' recommendations to cancel surgery. However, of the 24 written psychological reports, produced by six

Table 3. *Descriptive statistics for Short Form 20 (SF-20) Multi-Dimensional Health Survey sub-scales*

SF-20 sub-scale[a]	Psych. report mean (SD)	No psych. report mean (SD)	Normative mean (SD)[b]	Psych. report effect size	No psych. report effect size	Percentage in poor health[c]			
						Psych. report	No psych Report	Normative patients[b]	General population[d]
Physical functioning	43.4 (30.4)	49.0 (30.6)	78.5 (30.8)	−1.14	−0.96	93	91	45	22
Role functioning	43.1 (45.8)	51.5 (42.7)	77.5 (38.3)	−0.90	−0.68	66	63	28	12
Social functioning	71.7 (34.9)	72.7 (34.4)	87.2 (23.6)	−0.66	−0.61	24	27	9	[e]
Mental health index	67.3 (23.3)	62.4 (23.0)	72.6 (20.2)	−0.26	−0.50	41	55	31	19
Current health perceptions	56.0 (29.6)	54.4 (28.3)	63.0 (26.8)	−0.26	−0.32	58	67	52	20
Pain severity	57.9 (24.1)	55.7 (29.2)	31.4 (27.7)	0.96	0.89	76	62	29	[e]

[a] Observed range of all scores was 0–100. A high score indicates better health except for pain, where a high score indicates more pain.

[b] The normative sample consisted of patients presenting to physicians, psychologists and other mental health providers within HMOs, multi-specialty groups, and solo fee-for-service groups.

[c] Poor health defined as: physical and role functioning means one or more limitations; social functioning means limitations much of the time;mental health means lowest 19% of scores in general population (score of 67 or lower); health perceptions means the lowest 20% of scores in the general population (score of 70 or lower); pain means moderate, severe or very severe pain.

[d] General population refers to a random sample of subjects selected for telephone administration of the SF-20 (Stewart *et al.* 1988).

[e] Not available.

different psychologists, that we were able to locate, 20 recommended proceeding with surgery. In only four cases did the psychologist recommend against surgery but the surgeon proceeded anyway. The outcomes for these four patients did not differ from the overall psychological evaluation cohort. A second possibility is that psychologists recommended surgery along with specific psychological interventions to reduce the risk for those with potential complications, but these recommendations were either inadequate, ignored or both. Of the 20 reports that recommended surgery, ten made no recommendations for any pre- or post-surgical psychosocial intervention, three made recommendations for pre-surgical interventions only, six recommended post-surgical treatments only, and one suggested both. Typically these recommendations were quite non-specific in their instructions as to how to proceed. For example, the suggested pre-surgical interventions included 'brief supportive counselling', clarification of surgical expectations, counselling regarding 'the necessity to accept a certain amount of pain' and the need for 'evidence that the patient has solved his alcohol, tobacco, and pain medication problems'. Post-surgical interventions mainly focused on vocational counselling and the necessity for compliance with any physical therapy or rehabilitation regimen but generally lacked specific detail. We were not able to determine the extent to which recommendations were actually communicated to patients or how often patients followed-up on them. However, there are indirect data that bear on this question. There were no differences in money spent for rehabilitation services between those referred for psychological evaluation and those not referred and rehabilitation costs did not generally correlate with outcome measures.[1] There is no reason to believe that any of the recommendations were iatrogenic, however they may not have been particularly effective. Indeed, 50% of the patients were completely cleared for surgery with no recommendations at all. As noted, psychologists often provided recommendations in a general, and perhaps difficult to implement, form. These observations lead us to tentatively assert that pre-surgical psychological evaluations may be more effective if psychologists presented recommendations in more specific ways adapted to this specialized medical setting. Notably, recommendations should be expressed in precise behavioural terms that consider the particular clinical setting, resources available and feasability. These written recommendations should also be highlighted within the report and prioritized. Finally, recommendations should be communicated to surgeons in concise, clear, specific, unequivocal, and detailed ways (cf. Gregory, 1999; Haley *et al.*, 1998; Popkin & Mackenzie, 1984).

A third possibility to explain these results is that psychologists may not have the tools necessary to predict how psychological factors influence lumbar surgical outcomes. However, there is a research literature demonstrating that psychological variables influence surgical outcomes. Specifically, cognitive catastrophizing, perceived loss of control, litigation, job dissatisfaction, solicitous partners and marital dissatisfaction have all been identified (cf. Block, 1999; Block & Callewart, 1999). Further, numerous studies suggest the MMPI and MMPI-2 are useful in predicting lumbar surgery outcomes (Doxey *et al.*, 1988; Kuperman *et al.*, 1979; Masters *et al.*, 2000; Oostdam *et al.*, 1981; Riley *et al.*, 1995; Turner & Leiding, 1985; Wilfling *et al.*, 1973). Interestingly, nearly all of the psychologists in this study used the MMPI-2 in their assessment.

Given that the literature suggests that psychologists have tools available to make predictions about lumbar fusion outcomes, and the results of those in this study who underwent pre-surgical psychological evaluation were not optimal, one must wonder whether the empirically established evaluation methods are being properly applied in real world practice. Thus, it may be the methods used by psychologists fell short of what has been suggested in clinical research. Suggestions for improvement include using a standardized test battery, protocols for the clinical evaluation and decision-making algorithms. Readers inter-

ested in a new approach to surgical outcome prediction are referred to the Pre-surgical Psychological Screening Scale (Block, 2000; Block & Callewart, 1999).

It is noteworthy that surgeons correctly identified at least some individuals who were at increased risk for poor surgical outcome and referred them for psychological evaluation. What is unclear is the elucidation of the salient factors that play a major role in this decision. The only measured variables in the present study that distinguished those referred for evaluation from those not referred were the presence of depression and use of alcohol. Seemingly important physical, medical and historical variables were apparently not factors contributing to referral insofar as they did not discriminate referred patients from those not referred. Although depression and alcohol use may have been two of the characteristics important to the surgeons' discrimination process, many other patient factors were not studied and therefore their importance is unknown. For example, surgeons' perceptions of relevant patient personality characteristics (e.g. anxiety, somatization, malingering, pain magnification) might also be important in the referral decision. Additional study of the processes that surgeons use in ordering pre-surgical psychological evaluations is indicated.

Finally, this study again emphasizes the importance of using multiple measures of outcome. If arthrodesis rates were the only measure considered there would be no differences between the two groups. Arthrodesis is important, but in and of itself is not synonymous with successful surgical outcome. Payors are interested in expenses involved with the procedure, employers are interested in the likelihood of patients returning to work (and how long this will take), and patients are interested in pain reduction and functional improvement.

There are several limitations of this study. First, the work of only six psychologists practising in a specific geographical area and working with a compensated sample was considered. However, the psychologists in this study were experienced clinicians who remain well regarded for their work with these patients. This is a study of the current state of clinical practice. It does not address outcomes achieved if a 'best practice' model were used. The use of compensated patients always creates difficulties pertaining to motivation and other psychological variables that may not be present in non-compensated samples. But it is precisely these psychologically challenging patients for whom psychological evaluation is most needed. Additionally, there was no way to evaluate how and to what degree recommendations suggested by the psychologists were implemented. Many of the outcomes were not favourable for those receiving psychological evaluation, but the precise reasons are not clear.

This study raises many questions. For example, what results would emerge from a study comparing two groups of patients, both referred for psychological evaluation by their surgeon, but randomly assigned to actually receive the evaluation or not receive it? Are the types of pre- and post-surgical interventions typically recommended by psychologists, if competently applied, effective in improving surgical outcomes? Finally, can practising psychologists improve their surgical predictions?

Acknowledgements

This project was supported from grants provided by the University of Utah Medical School and the Labor Commission of Utah. Results presented in this paper should not be construed as an endorsement or indictment of lumbar fusion techniques for injured workers by the authors, the Labor Commission of Utah, Utah State University or the University of Utah. The authors wish to thank the valuable suggestions of an unknown reviewer of an earlier version of this manuscript.

Portions of this study were presented at the 11th Annual Convention of the American Psychological Society, Denver, Colorado, 1999.

Note

[1] Further details on these analyses are available from the principal author.

References

BLOCK, A.R. (1999). Presurgical psychological screening in chronic pain syndromes: psychosocial risk factors for poor surgical results. In: R. J. GATCHEL & D. C. TURK (Eds), *Psychosocial factors in pain: critical perspectives* (pp. 390–400). New York: Guilford Press.

BLOCK, A.R. (2000). Predictive efficacy of a pre-surgical psychological screening tool for back surgery patients. *Annals of Behavioral Medicine, 22*(Suppl.), S207.

BLOCK, A.R. & CALLEWART, C. (1999). Surgery for chronic spine pain: procedures for patient selection and outcome enhancement. In: A. R. BLOCK, E. F. KREMER & E. FERNANDEZ (Eds), *Handbook of pain syndromes: biopsychosocial perspectives* (pp. 191–212). Mahwah, NJ: Lawrence Erlbaum.

BOOS, N., MARCHESI, D. & AEBI, M. (1991). Treatment of spondylolysis and spondylolysthesis with Cotrel-Dubousset instrumentation: a preliminary report. *Journal of Spinal Disorders, 4,* 472–479.

BOOS, N., MARCHESI, D. & AEBI, M. (1992). Survivorship analysis of pedicular fixation systems in the treatment of degenerative disorders of the lumbar spine: a comparison of Cotrel-Dubousset instrumentation and the AO internal fixator. *Journal of Spinal Disorders, 5,* 403–409.

DEBERARD, M.S. (1998). Predicting lumbar fusion surgery outcomes from pre-surgical patient variables: the Utah lumbar fusion outcome study (Doctoral dissertation, Utah State University). *Dissertation Abstracts International, 59,* 4526.

DEBERARD, M.S., MASTERS, K.S., COLLEDGE, A., SCHLEUSENER, R. & SCHLEGEL, J. (2001). Outcomes of posterolateral lumbar fusion in Utah patients receiving workers' compensation: a retrospective-cohort study. *Spine, 26,* 738–747.

DEYO, R.A. (1986). Comparative validity of the Sickness Impact Profile and shorter scales for functional assessment in low-back pain. *Spine, 11,* 951–954.

DOXEY, N.C., DZIOBA, R.B., MITSON, G.L. & LACROIX, J.M. (1988). Predictors of outcome in back surgery candidates. *Journal of Clinical Psychology, 44,* 611–622.

FRANKLIN, G.M., HAUG, J., HEYER, N.J., MCKEEFREY, S.P. & PICCIANO, J.F. (1994). Outcome of lumbar fusion in Washington State workers' compensation. *Spine, 17,* 1897–1903.

GREGORY, R.J. (1999). *Foundations of intellectual assessment.* Boston: Allyn & Bacon.

HALEY, W.E., MCDANIEL, S.H., BRAY, J.H., FRANK, R.G., HELDRING, M., JOHNSON, S.B. *ET AL.* (1998). Psychological practice in primary care settings: practical tips for clinicians. *Professional Psychology: Research and Practice, 29,* 237–244.

KUPERMAN, S.K., OSMON, D., GOLDEN, C.J. & BLUME, H.G. (1979). Prediction of neurosurgical results by psychological evaluation. *Perceptual and Motor Skills, 48,* 311–315.

MASTERS, K.S., SHEARER, D.S. & OGLES, B.M. (2000). MMPI-2 cluster profiles predict one-year lumbar surgery outcomes. *Annals of Behavioral Medicine, 22* (Suppl.), S207.

OOSTDAM, E.M.M. & DUIVENVOORDEN, H.J. (1983). Predictablity of the result of surgical intervention in patients with low back pain. *Journal of Psychosomatic Research, 27,* 273–281.

OOSTDAM, E.M.M., DUIVENVOORDEN, H.J. & PONDAAG, W. (1981). Predictive value of some psychological tests on the outcome of surgical intervention in low back pain patients. *Journal of Psychosomatic Research, 25,* 227–235.

POPKIN, M.K. & MACKENZIE, T.B. (1984). Communicating with the referring physician. In: F. GUGGENHEIM & M. WEINER (Eds), *Manual of psychiatric consultation and emergency care* (pp. 115–123). New York: Jason Aronson, Inc.

RILEY, J.L., ROBINSON, M.E., GEISSER, M.E., WITTMER, V.T. & SMITH, A.G. (1995). Relationship between MMPI-2 cluster profiles and surgical outcome in low-back pain patients. *Journal of Spinal Disorders, 8,* 213–219.

ROLAND, M. & MORRIS, R. (1983a). 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, 8,* 141–144.

ROLAND, M. & MORRIS, R. (1983b). A study of the natural history of low back pain: Part II: Development of guidelines for trials of treatment in primary care. *Spine, 8,* 145–150.

STAUFFER, R. & COVENTRY, M.B. (1972). Anterior interbody lumbar spine fusion. *Journal of Bone and Joint Surgery, 54,* 756–768.

STEWART, A.L. & WARE, J.E. (Eds) (1992). *Measuring functioning and well being: the medical outcomes study approach.* Durham and London: Duke University Press.

STEWART, A.L., HAYS, R.D. & WARE, J.E. (1988). The MOS short-form general health survey. Reliability and validity in a patient population. *Medical Care, 26,* 724–735.

TAYLOR, M.E. (1989). Return to work following back surgery: a review. *American Journal of Industrial Medicine, 16,* 79–88.

TAYLOR, V.M., DEYO, R.A., CHERKIN, D.C. & KREUTER, W. (1994). Low back hospitalizations: recent US trends and regional variations. *Spine, 19,* 1207.

TURNER, J.A. & LEIDING, W.C. (1985). Correlation of the MMPI with lumbosacral spine fusion results: prospective study. *Spine, 10,* 932–936.

TURNER, J.A., ERSEK, M., HERRON, L., HASELKORN, J., KENT, D., CIOL, M.A. & DEYO, R. (1992). Patient outcomes after lumbar spinal fusions. *Journal of the American Medical Association, 268,* 907–911.

UOMOTO, J.M., TURNER, J.A. & HERRON, L.D. (1988). Use of the MMPI and MCMI in predicting outcome of lumbar laminectomy. *Journal of Clinical Psychology, 44,* 191–197.

WILFLING, F.J., KLONOFF, H. & KOKAN, P. (1973). Psychological, demographic and orthopaedic factors associated with prediction of outcome of spinal fusion. *Clinical Orthopaedics, 90,* 153–160.