

Emotional Health Predicts Pain and Function After Fusion: A Prospective Multicenter Study

Paula M. Trief, PhD,* Robert Ploutz-Snyder, PhD,† and Bruce E. Fredrickson, MD‡

Study Design.

Prospective.

Objectives. To assess whether patients with poorer emotional health before fusion surgery will have worse pain and function after surgery, and to identify patient variables that predict fusion outcomes.

Summary of Background Data. A systematic review of fusion outcomes studies reported an average of 68% “satisfactory” outcomes. The persistence of pain and functional limitations leads to emotional and financial costs. Therefore, it is important to identify the factors that affect fusion outcome. Research has explored psychosocial factors (e.g., depression, anxiety) as being important outcome predictors.

Methods. Data from subjects enrolled in a multisite trial of 2 fusion systems were analyzed. Subjects completed measures of health-related quality of life (SF-36), pain (visual analog scale), and function (Oswestry Disability Index). A total of 160 subjects completed measures before fusion, 155 completed measures 12 months after fusion, and 115 completed measures 24 months after fusion. Subject variables (i.e., age, gender, smoking, workers’ compensation, and second surgery status) and presurgical pain/function were regressed on pain/function outcomes after surgery. This model was compared to one that included presurgical Mental Component Scores (MCSs), which is a SF-36 derived measure of emotional health, to determine whether MCS data significantly improved the prediction of pain/function.

Results. Higher presurgical MCS (i.e., better emotional health) predicted less back and leg pain after surgery. Similarly, higher presurgical MCS predicted better physical function after surgery. Other important predictors of pain and function were presurgical pain and function, workers’ compensation, and smoking status. The associations were modest (2% to 9% of independent variance accounted for), but significant.

Conclusions. Presurgical emotional status is one significant predictor of pain and function outcomes up to 2 years after fusion. Other significant predictors included workers’ compensation status, smoking status, and presurgical pain/function. Studies to identify and intervene

with patients with poorer emotional status will clarify whether presurgical mental health intervention can improve pain and function outcomes after surgery or whether these patients are not candidates for surgery.

Key words: fusion, quality of life, depression, pain, function. **Spine 2006;31:823–830**

Back pain has become the leading cause of disability and lost production in the United States.¹ In 1997, direct and indirect costs associated with industrial-related back pain in the United States were estimated to exceed \$170 billion per year.² The National Center for Health Statistics estimated that more than 317,000 back surgeries were performed in 1997 to address this problem, including 192,000 spinal fusions.^{3,4} A systematic review of spinal fusion outcome studies from 1966 to 1991 reported an average of 68% “satisfactory” outcomes.⁵ The authors report a range of 16% to 95% success. This result reflects differences in the operational definitions of “satisfactory outcome” (e.g., “solid fusion”), patient satisfaction, or return to work. In addition, many of these studies included research of fusions that have a high medical failure rate, or pseudarthrosis, if bone graft alone^{6,7} or posterior pedicle screw fixation^{8,9} is used.

Interbody fusion devices, such as threaded cages, have been developed, and a recent review concludes that these devices have improved fusion success rates and subjective clinical effectiveness,¹⁰ although overuse and failures have been noted.^{11,12} In a 4-year follow-up study of 196 patients who underwent fusion using the Bagby and Kuslich lumbar fusion cage, technically successful fusion occurred in 98% of patients, and patient reports of pain and function showed “significant improvement.”¹³ Pain rating scale (1 = none to 6 = disabling pain) scores decreased from an average of 5.1 (± 0.65) to 3.4 (± 1.02), and function scores (7 activities of daily living scales) also improved, decreasing from 21 to 15, with the higher number being poorer function, and normal individuals scoring between 9 and 12.

Although significant, one sees that most patients continued to have mild-to-moderate residual pain and poorer than “normal” function. Similarly, Penta and Fraser¹⁴ reviewed 108 patients with fusion 10 years after surgery and found that 78% rated themselves as having had “complete relief” or “a good deal” of relief, indicating that 22% had poor results and that even the successes had persistent problems. The persistence of pain and function limitations leads to emotional and financial costs for patients, families, and society.

From the *Department of Psychiatry, †Center for Outcomes Research and Evaluation, and ‡Department of Orthopedic Surgery, State University of New York Upstate Medical University, Syracuse, NY.

Acknowledgment date: May 12, 2004. First revision date: July 19, 2004. Second revision date: November 8, 2004. Third revision date: February 16, 2005. Fourth revision date: April 1, 2005. Fifth revision date: May 3, 2005. Acceptance date: May 4, 2005.

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

Corporate/Industry funds were received in support of this work. One or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript: e.g., honoraria, gifts, consultancies.

Address correspondence and reprint requests to Paula M. Trief, PhD, Department of Psychiatry, State University of New York Upstate Medical University, 750 E. Adams Street, Syracuse, NY 13210; E-mail: triefp@upstate.edu

Given the data, it is important to identify the factors that may impact on fusion outcome and surgical selection, as well as appropriate measures to assess them. A body of literature explores psychosocial factors as important predictors of outcome.¹⁵ Most of this research has used the Minnesota Multiphasic Personality Inventory and its revision to identify “cluster profiles” that predict poor surgical outcome.^{16–20} However, critics have questioned the clinical use of this measure, and the fact that most of the data are retrospective and correlational.^{21–23} One prospective study found that “psychological disturbance,” measured by pain drawings and inappropriate signs, predicted fusion outcome.²⁴ Our recent prospective study of 102 patients who underwent lumbar spine surgery (mixed types) found that lower presurgical anxiety,²⁵ depression,²⁶ and hostility²⁷ predicted better outcome (*i.e.*, return to work, pain decrease, functional improvement).²⁸ However, the surgeries were heterogeneous, the measures used are not routinely part of orthopedic assessment, and the follow-up assessment point was only 1 year, a relatively short period of time.

The primary purpose of the current study was to determine whether presurgical emotional health, measured by a valid and reliable tool, predicts outcomes after surgery in a relatively homogeneous group (*i.e.*, all patients undergo lumbar fusion), both at short and long-term follow-up. Secondary aims were to explore whether other psychosocial factors predict outcomes after fusion. In the current study, we used the Medical Outcomes Study-Short Form (SF-36), a measure that is fast becoming the gold standard for assessing physical and emotional health outcomes.²⁹ The institutional review board of State University of New York Upstate Medical University and the institutional review boards of other participating institutions in which data were collected approved the study.

■ Materials and Methods

Subjects. Data from participants in a larger clinical trial were analyzed. This was a Food and Drug Administration approved multicenter (10 sites) investigational device experiment, in which patients undergoing lumbar fusions were randomly assigned to either BAK or InFix lumbar cage conditions. All patients underwent fusion performed at the same sites and by the same physicians; the only difference was the type of cage used. A study coordinator at each site recruited patients when they presented at physicians’ offices. The study coordinator screened for inclusion/exclusion criteria established for the trial. Individuals were included if they had pain, functional deficit, or neurologic deficit for more than 6 months and were diagnosed with degenerative disc disease. Degenerative disc disease was defined according to specific criteria, including low back pain of a discogenic nature, confirmed by patient symptoms and history, physical examination findings, and associated abnormal findings on magnetic resonance imaging or other radiologic evaluations that led the treating physician to identify the disc as the source of the pain.

Subjects were between 21 and 70 years of age, had a maximum of 2 diseased levels to be instrumented, had undergone no more than 2 previous nonfusion surgeries at the same level, and had failed continued conservative treatment for a minimum of 6 months. Patients were excluded per the larger trial criteria if they had ≥ 2 of the following psychosocial factors: alcoholic/drug dependent or treated for alcoholism/drug dependency within past 5 years, current or past spinal surgery litigation, and smoking more than 1 pack per day (assessed by self-report). Subjects were educated about anterior lumbar interbody fusion and the requirements of the study. On signing the approved consent document, a blinded randomization envelope was opened to determine treatment group. Subjects in this trial completed measures of pain, function, and health-related quality of life. These questionnaires were completed at a presurgical visit to the clinic no more than 1 month before surgery (baseline), and at clinic follow-up visits 1 and 2 years after fusion surgery.

Measures

Medical Outcomes Study-Short Form (SF-36). The SF-36 is one of the most widely used questionnaires in health research used to assess health-related quality of life.²⁹ It has excellent reliability and validity.^{30–32} It contains 36 multiple-choice questions, scored on 8 subscales: physical function, role limitations caused by physical problems, bodily pain, general health, vitality, social functioning, role limitations caused by emotional problems, and mental health. Scores on the subscales can be aggregated to create a Physical Component Score (PCS) and Mental Component Score (MCS).³³ The MCS is calculated using scores on the social functioning, role limitations caused by emotional problems, mental health, and vitality subscales.

Visual Analog Scale (VAS). As is the standard, presenting subjects with a 100-mm horizontal line and asking them to put an X on the line that represented the level of pain during the past week is used to obtain VAS scores. The range was 0 = “no pain at all” to 100 = “worst pain imaginable.”³⁴ The VAS was used to assess both back and leg pain.

Oswestry Disability Index (modified). The Oswestry is a commonly used measure of functional status related to limitations associated with low back pain. It is comprised of 10 questions that describe back-related disability as a combination of physical and social restriction in various aspects of daily life. A sum is calculated and presented as a percentage (range 0% = no disability to 100% = the worst possible disability).³⁵ It has been reliable and valid as a measure of functional disability.^{36,37} Because of concerns about the intrusiveness and potential discomfort elicited by a question concerning sexuality, researchers in the larger trial chose to omit it, thus changing the possible range (0–90), but not changing the percentage of disability.

Statistical Analysis. Blocked multiple regression analyses were used to evaluate the combined and independent variance contributions of various predictors of pain and function following surgery. Specifically, presurgical back pain, height, weight, age, gender, smoker status (yes, no), workers’ compensation status (yes, no), and whether patient underwent a second surgery (yes, no) were entered into the initial regression predicting 12-month pain. A total of 22 subjects (13.8%) underwent a second surgical procedure after the index surgery, but before the follow-up assessments, generally because of medical

complications. This was labeled “underwent second surgery.” Thus, for these subjects, outcomes represent results after 2 surgeries, and this variable was also entered into the first block.

We entered participants’ presurgical MCS data into the second block to determine whether MCS was a significant predictor of 12-month back pain after factoring in the initial block of predictors. We conducted similar analyses predicting 12-month function (SF-36 PCS, Oswestry), and 24-month pain and function. In each case, we entered baseline pain or function scores into the model to control for presurgical pain or function. Participants were eliminated from any regression analysis if they did not supply valid data for all predictors and outcomes (*i.e.*, list-wise elimination). Therefore, because each regression uses only subjects that have valid data for all predictors and the relevant outcomes (*i.e.*, completers), each regression used a slightly different sample, and the number of subjects varied slightly.

In all models, we interpret the multiple R^2 and associated P values as an evaluation of the overall model, and we compared the reduced (base) model to the full model, including MCS, by comparing the amount of unexplained variance in either model using the analysis of variance statistic. We provide the standardized β coefficients and associated P values to assess statistical significance of individual predictors, and we evaluate the independent variance contributions of predictors by examination of their squared semi-partial correlation (sr^2), representing the amount of a reduction in the overall R^2 the model would sustain if a predictor is eliminated from the model.

Also, to begin to explore questions of clinical significance, we were interested in assessing how much of an improvement in preoperative MCS is associated with a clinically meaningful difference in function and back pain. Preliminary work by Lurie *et al*³⁸ has suggested that a 5-point increase on the SF-36 PCS, and a 7-point increase in Oswestry scores, represent a clinically meaningful improvement in function, however, they only followed subjects for 3 months. We used the unstandardized β coefficients from the regression models to project pain and function outcomes (12 and 24-month) for the typical subject in this trial, and computed the average improvement in baseline MCS that is necessary to see an average increase of 5 points in PCS, and 7 points in Oswestry scores, and, thus, may be clinically meaningful.

■ Results

Subject Characteristics

We were provided with a data set of 171 subjects of the larger trial, which included all subjects who were enrolled in the trial at that time and, therefore, had agreed to be randomized to 1 of the 2 fusion cage systems, had undergone surgery at least 2 years prior, and thus could provide 24-month assessments. Although sites were requested to document the number of patients who underwent fusions and were approached to participate in this study but refused, the sponsor has not collected that data, and, thus, we cannot compare enrollees to refusers on baseline demographic data to ensure comparability. Of the 171 subjects enrolled in the fusion study, 6 did not complete baseline demographic data forms, and 5 did not complete the SF-36, despite agreeing to do so. These 11 individuals were eliminated from our analyses. This

Table 1. Demographics and Outcome Measures for 160 Subjects

	No. Subjects	Mean (SD)	%
Demographic/predictor variables			
Male gender	160		51.9
Workers’ compensation–Yes	160		37.5
Current smoker	160		45
Second surgery needed	160		13.8
Age (ys)	160	44.22 (8.59)	
Height (in)	160	67.39 (6.67)	
Weight (lbs)	160	190.19 (41.46)	
Presurgical MCS	160	41.54 (13.00)	
Presurgical back pain	158	72.96 (21.10)	
Presurgical leg pain	158	61.26 (27.53)	
Presurgical Oswestry function	160	60.19 (17.16)	
Presurgical SF-36 PCS	160	28.41 (6.27)	
Outcome measures			
VAS back pain 12 mos	151	43.85 (30.47)	
VAS back pain 24 mos	115	44.84 (31.89)	
VAS leg pain 12 mos	148	35.49 (31.56)	
VAS leg pain 24 mos	115	38.12 (31.91)	
Oswestry –12 mos	153	38.86 (25.48)	
Oswestry –24 mos	115	40.00 (26.28)	
SF-36 PCS –12 mos	151	36.60 (11.02)	
SF-36 PCS –24 mos	114	36.13 (11.97)	

SD indicates standard deviation.

resulted in 160 subjects who completed the baseline assessment and had undergone surgery for our analyses (Table 1). Because of the small number ($n = 11$) of subjects eliminated, we could not statistically compare them to our larger sample.

There were 83 men (51.9%) and 77 women (48.1%), with an age range of 26–67 years (mean 44.22, standard deviation 8.59). A total of 60 subjects (37.5%) were receiving workers’ compensation disability payments, and 72 (45%) were current smokers. All 160 subjects provided valid pain or function outcomes data for at least 1 of our follow-up assessments (12 or 24 months). However, some outcomes data are missing for a few

Table 2. Pain and Function Change From Presurgical Baseline

Outcome Measure	No. Subjects	Mean (SD)	Change From Baseline*
Baseline back pain (VAS)	105	74.77 (21.49)	
12-month back pain	105	45.30 (31.53)	$P < 0.001$
24-month back pain	105	44.53 (32.02)	$P < 0.001$
Baseline leg pain (VAS)	103	61.27 (27.85)	
12-month leg pain	103	37.07 (32.34)	$P < 0.001$
24-month leg pain	103	38.43 (31.99)	$P < 0.001$
Baseline Oswestry	109	60.62 (16.25)	
12-month Oswestry	109	38.19 (26.01)	$P < 0.001$
24-month Oswestry	109	39.75 (26.20)	$P < 0.001$
Baseline SF-36 PCS	109	28.51 (6.14)	
12-month SF-36 PCS	109	36.83 (11.42)	$P < 0.001$
24-month SF-36 PCS	109	36.32 (12.13)	$P < 0.001$

*Repeated measures analysis of variance with a priori contrasts comparing each postoperative measure to baseline. The number of subjects for whom preoperative, and 12 and 24-month data were available for statistical comparison determines sample sizes (No.) in this table.

SD indicates standard deviation.

Table 3. Prospective Multiple Regression Analyses Predicting Postsurgical Function (2 measures) From Baseline Characteristics and Significance of Difference Between Models With and Without Presurgical MCS Data

Model*	β †	<i>P</i>	sr ²	Multiple R ²	Model Significance	Full vs. Reduced <i>P</i> Change‡
Predicting 12-month function (SF-36 PCS) (n = 151)						
Age (ys)	-0.09	0.24	0.01§	0.25	<0.001	<0.01
Gender (1 male, 2 females)	0.02	0.85	0.00			
Height (in)	0.01	0.88	0.00			
Weight (lbs)	-0.02	0.78	0.00			
Workers' compensation (0 = no, 1 = yes)	-0.20	0.01	0.03			
Current smoker (0 = no, 1 = yes)	-0.17	0.03	0.03			
Second surgery needed (0 = no, 1 = yes)	-0.18	0.02	0.03			
Presurgical function (PCS)¶	0.32	<0.01	0.08			
Presurgical MCS (MCS)	0.22	0.01	0.04			
Predicting 24-month function (SF-36 PCS) (n = 114)						
Age	-0.08	0.64	0.01	0.35	<0.001	<0.001
Gender	0.13	0.21	0.01			
Height	0.19	0.11	0.03			
Weight	-0.06	0.57	0.00			
Workers' compensation	-0.16	0.07	0.02			
Current smoker	-0.13	0.14	0.01			
Second surgery needed	-0.17	0.04	0.03			
Presurgical function (PCS)¶	0.38	<0.01	0.10			
Presurgical MCS¶	0.35	<0.01	0.09			
Predicting 12-month function (Oswestry) (n = 153)						
Age	0.05	0.53	0.00	0.31	<0.001	0.05
Gender	-0.05	0.52	0.00			
Height	-0.07	0.34	0.00			
Weight	-0.04	0.62	0.00			
Workers' compensation	0.20	0.01	0.04			
Current smoker	0.19	0.01	0.03			
Second surgery needed	0.16	0.03	0.02			
Presurgical Oswestry¶	0.30	<0.01	0.07			
Presurgical MCS	-0.15	0.05	0.02			
Predicting 24-month function (Oswestry) (n = 115)						
Age	0.11	0.22	0.01	0.26	<0.001	0.10
Gender	-0.22	0.06	0.03			
Height	-0.25	0.05	0.04			
Weight	-0.02	0.85	0.00			
Workers' compensation	0.20	0.04	0.03			
Current smoker	0.12	0.19	0.01			
Second surgery needed	0.11	0.18	0.01			
Presurgical Oswestry	0.23	0.02	0.04			
Presurgical MCS	-0.16	0.10	0.02			

*The β coefficients, *P* values, sr² values, and multiple R² values were calculated on the full model that includes all predictors shown.

†The β coefficients are standardizations of the raw β coefficient. They are informative about relative contributions of predictors because the unit of measurement has been standardized.

‡The full versus reduced *P*-change column indicates the significance of comparing this model to a reduced model that included all predictors except presurgical MCS using hierarchical regression. Thus, *P*-change statistics indicate the extent to which the additional predictor (MCS) in the full model added predictive variance to the base model.

§Squared semi-partial correlation coefficients (sr²) are a measure of the independent predictive value of each predictor in the model. It represents the amount of variance that the overall R² would decrease if the predictor were eliminated from the model.

||Significant contribution at *P* < 0.05.

¶Significant contribution at *P* < 0.01.

subjects, thus, the sample sizes for our 8 regression models that predict pain and function vary. Specifically, the number of subjects who completed the 12 and 24-month VAS back pain scales was 151 (94%) and 115 (97%), respectively, and who completed the VAS leg pain scales was 148 (93%) and 115 (97%), respectively (Table 2).

The number of subjects who completed 12 and 24-month Oswestry scales was 153 (96%) and 115 (72%), respectively, and 12 and 24-month PCS scales was 151 (94%) and 114 (71%), respectively (Table 3). The Student *t* test compared baseline MCS, and measures of pain and function of subjects who provided all predictors and outcomes for regression analysis (*i.e.*, completers) to subjects who were eliminated from that analysis because they were

missing the relevant 12 or 24-month outcome data (*i.e.*, noncompleters). We found no statistically significant differences between these groups on baseline pain or function (*P* values ranged from 0.07 to 0.78), but 12-month completers had, on average, higher MCSs (*P* < 0.004).

Overall Surgical Outcome Results

Table 4 presents data on baseline and follow-up pain and function outcomes. Repeated measures analyses of variance with a priori contrasts, comparing baseline to each follow-up, show highly significant improvements in back pain, leg pain, and reported function at both 12 and 24-month follow-up.

Table 4. Prospective Multiple Regression Analyses Predicting Postsurgical Back/Leg Pain From Baseline Characteristics and Significance of Difference Between Models With and Without Presurgical MCS Data

Model*	β †	<i>P</i>	sr ²	Multiple R ²	Model Significance	Full vs. Reduced <i>P</i> Change‡
Predicting 12-month back pain (n = 149)*				0.29	<0.001	<0.01
Age (ys)	0.04	0.64	0.00§			
Gender (1 male, 2 females)	-0.06	0.47	0.00			
Height (in)	0.05	0.57	0.00			
Weight (lbs)	-0.04	0.60	0.00			
Workers' compensation (0 = no, 1 = yes)	0.12	0.11	0.01			
Current smoker (0 = no, 1 = yes)	0.18	0.02	0.03			
Second surgery needed (0 = no, 1 = yes)	0.21	<0.01	0.04			
Presurgical back pain (PCS)	0.27	<0.01	0.07			
Presurgical MCS (MCS)	-0.24	<0.01	0.05			
Predicting 24-month back pain (n = 113)				0.24	0.001	<0.01
Age	0.09	0.35	0.00			
Gender	-0.26	0.03	0.04			
Height	-0.19	0.14	0.02			
Weight	0.08	0.45	0.00			
Workers' compensation	0.19	0.04	0.03			
Current smoker	0.16	0.10	0.02			
Second surgery needed	0.14	0.11	0.02			
Presurgical back pain	0.12	0.18	0.01			
Presurgical MCS	-0.26	<0.01	0.06			
Predicting 12-month leg pain (n = 146)				0.30	<0.001	0.03
Age	0.17	0.03	0.03			
Gender	0.03	0.69	0.00			
Height	-0.08	0.34	0.00			
Weight	0.07	0.42	0.00			
Workers' compensation	0.27	<0.01	0.06			
Current smoker	0.24	<0.01	0.05			
Second surgery needed	0.10	0.18	0.01			
Presurgical leg pain	0.21	0.01	0.04			
Presurgical MCS	-0.17	0.03	0.03			
Predicting 24-month leg pain (n = 113)				0.31	<0.001	0.03
Age	0.17	0.06	0.02			
Gender	-0.13	0.24	0.01			
Height	-0.30	0.02	0.04			
Weight	0.13	0.21	0.01			
Workers' compensation	0.26	<0.01	0.06			
Current smoker	0.12	0.18	0.01			
Second surgery needed	0.00	0.99	0.00			
Presurgical leg pain	0.20	0.02	0.04			
Presurgical MCS	-0.20	0.03	0.03			

*The β coefficients, *P* values, sr² values, and multiple R² values were calculated on the full model that includes all predictors shown.

†The β coefficients are standardizations of the raw β coefficient. They are informative about relative contributions of predictors because the unit of measurement has been standardized.

‡The full versus reduced *P*-change column indicates the significance of comparing this model to a reduced model that included all predictors except presurgical MCS using hierarchical regression. Thus, *P*-change statistics indicate the extent to which the additional predictor (MCS) in the full model added predictive variance to the base model.

§Squared semi-partial correlation coefficients (sr²) are a measure of the independent predictive value of each predictor in the model. It represents the amount of variance that the overall R² would decrease if the predictor were eliminated from the model.

||Significant contribution at *P* < 0.05.

¶Significant contribution at *P* < 0.01.

Function

Our predictions of physical function (SF-36 PCS) indicate that higher presurgical MCS predicted better physical function at 12 ($\beta = 0.22$, *P* < 0.05) and 24 months ($\beta = 0.35$, *P* < 0.01) (Table 3). Smoking, workers' compensation, and second surgery were negatively related, and presurgical PCS was positively related to 12-month PCS function. Predictions of 24-month PCS revealed that only second surgery and presurgical PCS were additional predictors of PCS 2 years after surgery. Similar results were obtained using the Oswestry Disability Index (*i.e.*, presurgical MCS [$\beta = -0.15$, *P* < 0.05], smoking, workers' compensation status, and presurgical Oswestry scores all contributed significantly to the prediction of 12-month

Oswestry scores). Similarly, workers' compensation status, height, and presurgical Oswestry significantly predicted 24-month Oswestry scores, but there was only a trend for presurgical MCS to do so.

We also ran some exploratory analyses to address questions of clinical significance. As noted, preliminary work³⁸ suggests that a 5-point increase in SF-36 PCS represents a clinically meaningful change. There are limitations in basing our analyses on these data because our sample is more homogeneous, and our follow-up is longer. However, without other cutoff scores identified in the literature, we believed that it was a worthwhile exercise to ask the question of how much of a change in MCS would be associated with a clinically significant

improvement in function. We projected the effect of changing presurgical MCS on functional outcomes by creating a projection model that used the unstandardized regression coefficients from our models and hypothetical values for all other predictors in the model that are consistent with our “typical” subject.

Using the unstandardized regression coefficients from our regressions on 12 and 24-month function, the typical surgical patient would have a clinically meaningful change in 24-month PCS function (*i.e.*, at least a 5-point increase) if the preoperative MCS was 16 points higher. This result also generated a 3-point PCS improvement at 12 months. To achieve a clinically important change in 24-month Oswestry (*i.e.*, 7-point increase), a 23-point increase in preoperative MCS is statistically necessary; this MCS increase projects a 6.7-point improvement in 12-month Oswestry. These function change projections assume that all other predictors in the model remain constant, and, thus, we believe them to be somewhat conservative estimates because MCS is correlated with other predictors in the model, and those predictors also impact functional improvement.

Pain

Higher presurgical MCS (*i.e.*, better emotional health) predicted less back pain 12 months after surgery ($\beta = -0.24$, $P < 0.01$) (Table 4). Smoking, second surgery, and presurgical back pain were negatively related to VAS back pain ratings at 12 months. Presurgical MCS ($\beta = -0.26$, $P < 0.01$), gender, and workers' compensation status were also significant predictors of 24-month back pain. Similar results were found in predictions of leg pain, in which presurgical MCS was significantly predictive of leg pain at 12 ($\beta = -0.17$, $P < 0.05$) and at 24 months ($\beta = -0.20$, $P < 0.05$). Also, older age, receiving workers' compensation, being a smoker, and more presurgical leg pain were significantly related to more leg pain at 12 months. Significant predictors of 24-month leg pain include height, workers' compensation status, and presurgical leg pain.

The unique variance contribution (sr^2) of MCS data to 12-month back pain was 5%, second only to presurgical back pain data (7%). Similarly, the highest unique variance contribution to 24-month back pain was presurgical MCS (6%).

Again, we tried to explore the issue of clinical significance. Given that the literature provides no estimates of how much of a change in pain (as measured by the 100-mm VAS) would represent a clinically meaningful improvement, we projected 12 and 24-month pain improvements associated with a 16-point increase in preoperative MCS. We chose a 16-point MCS increase because this was previously suggested as being related to a clinically meaningful improvement in SF-36 PCS function. Our projections of pain showed that a 16-point increase in preoperative MCS would predict a change of 8.8 (12-month) and 10.1 (24-month) points on the VAS

back pain scale, and 6.6 (12-month) and 7.6 (24-month) points on the VAS leg pain scale.

Discussion

Anterior lumbar fusion is an increasingly commonly used surgical procedure to address chronic and persistent back and lower extremity pain caused by degenerative disc disease. However, persistent pain and dysfunction after fusion are frequently reported, and result in significant emotional and quality of life costs to the patient, as well as health care and legal monetary costs to society. Thus, predictors of outcome are needed.

In the current prospective study, we examined a specific surgery, anterior lumbar fusion with cages, and used a measure of emotional health that is commonly used in orthopedic offices, the SF-36, including a 2-year follow-up. Results indicate that patients who report poorer emotional status before fusion are significantly more likely to report poorer function outcomes 1 year after surgery and poorer pain outcomes a full 2 years after surgery. There were other variables that were of interest that were each significant predictors of outcome, at least at one time. Not surprisingly, those subjects who had more pain or poorer function before surgery also reported more pain or poorer function after surgery. Presurgical pain and function accounted for 2% to 10% of the variances, depending on outcomes and time of follow-up. However, presurgical back pain did not predict back pain after surgery at 24 months.

It is also not surprising that patients who underwent a second surgical procedure had poorer outcomes at follow-up, although this variable only accounted for 2% to 3% of variances. Smoking significantly predicted 12-month (range 2% to 3%), but not 24-month, pain and function. Workers' compensation status was a significant predictor (range 3% to 4%) in most analyses, yet emotional status was responsible for a higher percentage of the variance, and was generally the strongest independent predictor of back (range 5% to 6%) and leg pain (3%). Although presurgical function (range 4% to 10%) appears to be more important than emotional health (range 2% to 9%) when predicting function after surgery, this is not much of a difference. These percentages only represent the amount of independent variance, a conservative estimate, because it does not include the effect of emotional status that might be shared with other variables.

The overall outcomes of fusion surgery were very positive, at least for the subgroup of patients who completed the follow-up assessments. The common regression to the mean might have contributed to these findings. Also, the study is limited to results of fusion surgery performed with cages, and these results may not generalize to outcomes of other surgical procedures. We are also lacking some demographic and medical data (*e.g.*, race/ethnicity, education, medical comorbidities), which might have affected the results and limit generalizability. We were unable to obtain data on patients who

were recruited for the study but did not enroll and, therefore, do not know if this sample is truly representative of patients with fusion. Our sample may have had better (or worse) presurgical emotional status, and this difference could have affected results. We did find that for several of the analyses (12-month back pain, leg pain, and PCS function), the individuals who did not complete all assessments (*i.e.*, noncompleters) had poorer presurgical mental health (MCS) than completers. However, this result could strengthen our findings because this means that poorer mental health predicted poorer 12-month outcomes when individuals with even worse mental health were omitted from analyses because they dropped out of the study.

We performed additional preliminary data analyses to attempt to shed light on the “clinical significance” of the findings, and to understand how much improvement in presurgical mental health would be related to a clinically meaningful improvement in pain and function. To our knowledge, there are no accepted definitions or cutoff scores for what is deemed clinically significant with regard to chronic pain or function scores. Therefore, this aspect of the study was very exploratory and based on limited data (*i.e.*, unpublished, 3-month follow-up). We offer them to promote discussion of this issue. Significant improvement in presurgical MCS would be needed to achieve a meaningful change in postsurgical pain and PCS/Oswestry scores, which is not surprising because, although statistically significant, presurgical MCS only accounted for 2% to 9% of the independent variance in outcomes after surgery. The other significant predictors (*e.g.*, smoking, workers’ compensation status) did not independently explain a large percentage of the variance in outcomes. Further research is necessary to define clinically meaningful change, establish an appropriate time line to measure it, and more fully describe the psychosocial issues involved in achieving clinically meaningful change.

The finding in this prospective study that presurgical emotional health predicts function outcomes 1 year after surgery and pain outcomes a full 2 years after fusion is consistent with literature cited earlier. These data have been used to screen out high risk individuals. We believe that it is now time to develop interventions that help these individuals, rather than using this information to withhold surgery. Thus, we see 2 major challenges ahead. One is to develop evidence-based interventions that provide care for individuals who are recommended fusions, who are also having symptoms of poor mental health before surgery. A second challenge is for researchers to perform rigorous trials to determine whether these interventions result in improved surgical outcomes. We promote use of the SF-36 because it has high face validity, is easy to use, has become the gold standard for health-related quality of life assessment, and, thus, lends itself to translation into actual clinical practice.

Developing an appropriate and efficacious intervention will entail understanding the meaning of low MCS.

MCS less than 50 correlate with valid and reliable measures of mild major depression, moderate anxiety, low social support, and ineffective pain coping strategies.³⁹ Thus, effective interventions may need to improve coping skills to decrease depressive symptoms, cope with anxiety, and enhance social support.

A specific target may be depression. There is a growing body of literature that examines the relationship between pain and depression.⁴⁰ The term “depression-pain syndrome”⁴¹ has been coined to refer to the comorbidity noted between pain and depression, with evidence that having pain (or depression) appears to make depression (or pain) worse⁴² and that depression correlates with poor outcomes.^{43,44} Depression and pain may share common dysregulation of neurotransmitters, and, thus, both might be treated with antidepressants.^{45,46} Although there is strong evidence that cognitive behavioral therapy can improve depression,⁴⁷ the studies on treating depression and pain have most often focused on antidepressant medications, with mixed results reported.⁴⁰ In 1 study that explored this issue for patients with primary care arthritis, individuals with comorbid depression and arthritis were involved in a year-long depression treatment intervention, including medication plus a 6–8 session problem-solving psychotherapy program.⁴⁸ Intervention subjects, as compared to usual care, not only reported significant improvement in depressive symptoms, but also decreased pain, functional impairment, and arthritis-related daily activity interference, even though pain and arthritis were not addressed in the intervention. In contrast, another study examined cognitive-behavioral, pharmacologic, and no treatment of depression for individuals with rheumatoid arthritis.⁴⁹ This study found that only the antidepressant medication group improved and only in psychologic status; there was no improvement in pain or disease parameters. We agree with others who propose that a significant challenge is to develop cognitive-behavioral interventions that address the comorbidity issues, using an intervention model that combines assessment and treatment of both depression and pain.^{40,50}

■ Conclusions

Because fusion surgery may result in persistent pain and functional limitations, it is important to identify predictors of poor outcome. Our study supports the hypothesis that presurgical emotional health predicts long-term pain and function outcomes of fusion. Future studies should evaluate whether an intervention to improve emotional health before surgery can result in improved pain and function outcomes.

■ Key Points

- Many patients have persistent pain and functional limitations after fusion surgery.

- Presurgical emotional health predicts pain and function outcomes of fusion.
- Future studies should evaluate whether an intervention to improve emotional health before surgery can result in improved pain and function outcomes.

References

- Loeser JD, Bigos SJ, Muller G. Low back pain. In: Loeser JD, Butler SH, Chapman CR, et al., eds. *Bonica's Management of Pain*. 3rd ed. Philadelphia, PA: Lippincott, Williams & Wilkins; 2001:1508–28.
- Leigh JP, Markowitz SB, Fahs M, et al. Occupational injury and illness in the United States: Estimates of costs, morbidity, and mortality. *Arch Intern Med* 1997;157:68.
- National Center for Health Statistics. 1997 *National Hospital Survey*. Washington, DC: U.S. Department of Health and Human Services, Centers for Disease Control; 1997. Series 13, No. 144.
- Benson V, Marano MA. Current estimates from the National Health Interview Survey, 1995. National Center for Health Statistics. *Vital Health Stat* 1998;199:1–428.
- Turner JA, Ersek M, Herron L, et al. Patient outcomes after lumbar spinal fusions. *JAMA* 1992;268:907–911.
- Crock HV. Observations on the management of failed spinal operations. *J Bone Joint Surg Br* 1976;58:193–9.
- Stauffer RN, Coventry MB. Anterior interbody lumbar spine fusion: Analysis of Mayo Clinic series. *J Bone Joint Surg Am* 1972;54:56–68.
- Matzuzaki H, Tokyhashi Y, Matsumoto F, et al. Problems and solutions of pedicle screw plate fixation of lumbar spine. *Spine* 1990;15:1159–65.
- Ohlin A, Karlsson M, Duppe H, et al. Complications after transpedicular stabilization of the spine: A survivorship analysis of 163 cases. *Spine* 1994;19:2774–9.
- McAfee PC. Interbody fusion cages in reconstructive operations on the spine. *J Bone Joint Surg Am* 1999;81:859–80.
- O'Dowd JK, Lam K, Mulholland RC, et al. BAK cage: Nottingham results. Paper presented at: Annual Meeting of the North American Spine Society; 1998; San Francisco, CA.
- Onesti ST, Ashkenazi E. The Ray threaded fusion cage for posterior lumbar interbody fusion. *Neurosurgery* 1998;42:200–5.
- Kuslich SD, Danielson G, Dowdle JD, et al. Four-year follow-up results of lumbar spine arthrodesis using the Bagby and Kuslich lumbar fusion cage. *Spine* 2000;25:2656–62.
- Penta M, Fraser RD. Anterior lumbar interbody fusion: A minimum of 10 year follow-up. *Spine* 1997;22:2429–34.
- Block AR. *Pre-surgical Psychological Screening in Chronic Pain Syndromes*. Mahwah, NJ: Lawrence Erlbaum Associates; 1996.
- Doxey NCS, Dzioba RB, Mitson GL, et al. Predictors of outcome in back surgery candidates. *J Clin Psychol* 1988;44:611–22.
- Riley JI, Robinson ME, Geisser ME, et al. Relationship between MMPI-2 cluster profiles and surgical outcome in low-back pain patients. *J Spinal Disord* 1995;8:213–9.
- Spengler DM, Ouelette EA, Battie M, et al. Elective discectomy for herniation of a lumbar disc. *J Bone Joint Surg Am* 1990;12:230–7.
- Herron LD, Turner J, Clancy S, et al. The differential utility of the MMPI: A predictor of outcome in lumbar laminectomy for disc herniation vs. spinal stenosis. *Spine* 1986;11:847–50.
- Masters KS, Shearer DS, Ogles BM, et al. Presurgical MMPI-2 cluster profiles predict 1-year low-back surgery outcomes. *J Clin Psychol Med Settings* 2003;10:279–88.
- Helmes E, Reddon JR. A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychol Bull* 1993;113:453–71.
- Chapman SI, Pemberton JS. Prediction of treatment outcome from clinically derived MMPI clusters in rehabilitation for chronic low back pain. *Clin J Pain* 1994;10:267–76.
- Turk DC, Fernandez E. Personality assessment and the Minnesota Multiphasic Personality Inventory in chronic pain: Underdeveloped and overexposed. *PRN Forum* 1995;4:104–7.
- Greenough CG, Taylor LJ, Fraser RD. Anterior lumbar fusion: Results, assessment techniques and prognostic factors. *Eur Spine J* 1994;3:225–30.
- Spielberger CD, Gorsuch RL, Lushere RE. *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press; 1983.
- Zung WWK. A self-rating depression scale. *Arch Gen Psychiatry* 1965;12:63–70.
- Cook WW, Medley DM. Proposed hostility and Pharisac-virtue scales for the MMPI. *J Appl Psychol* 1954;38:414–28.
- Trief PM, Grant W, Fredrickson B. A prospective study of psychological predictors of lumbar surgery outcome. *Spine* 2000;25:2616–21.
- Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
- Ware JE, Snow KK, Kosinski M, et al. *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: New England Medical Center, The Health Institute; 1993.
- McHorney CA, Ware JE, Raczek AE. The MOS 36-item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;31:247–63.
- McHorney CA, Ware JE, Lu JFR, et al. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions and reliability across diverse patient groups. *Med Care* 1994;32:40–66.
- Ware J, Kosinski M, Keller SD. *SF-36 Physical and Mental Component Summary Scales: A User's Manual*. Boston, MA: The Health Institute, New England Medical Center; 1994.
- McCormack HM, Horne DJ, Sheather S. Clinical applications of visual analogue scales: A critical review. *Psychol Med* 1988;18:1007–19.
- Fairbank JC, Couper J, Davies JB, et al. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy* 1980;66:271–3.
- Baker DJ, Pynsent PB, Fairbank J. The Oswestry Disability Index revisited: Its reliability, repeatability and validity, and a comparison with the St. Thomas's Disability Index. In: Roland M, Jenner JR, eds. *Back Pain: New Approaches to Rehabilitation and Education*. Manchester, UK: Manchester University Press; 1989:174–86.
- Leclaire R, Blier F, Fortin L, et al. A cross-sectional study comparing the Oswestry and Roland-Morris Functional Disability Scales in two populations of patients with low back pain of different levels of severity. *Spine* 1997;22:68–73.
- Lurie JD, Hanscom B, Lee L, et al. Outcome measures in spine patients: What is a clinically important change? Paper presented at: 2001 Meeting of the International Society for the Study of the Lumbar Spine; June 19–23, 2001; Edinburgh, Scotland.
- Ayers DC, Franklin PD, Trief PM, et al. Psychological attributes of preoperative total joint replacement patients: Implications for optimal physical outcome. *J Arthroplasty* 2004;19(suppl 2):125–30.
- Bair MJ, Robinson RL, Katon W, et al. Depression and pain comorbidity: A literature review. *Arch Intern Med* 2003;163:2433–4.
- Lindsay PG, Wyckoff M. The depression-pain syndrome and its response to antidepressants. *Psychosomatics* 1981;22:571–3.
- Gallagher RM, Verma S. Managing pain and co-morbid depression: A public health challenge. *Semin Clin Neuropsychiatry* 1999;4:203–20.
- Painter JR, Seres JL, Newman RI. Assessing benefits of the pain center; why some patients regress. *Pain* 1980;8:101–13.
- Cherkin DC, Deyo R, Street JH, et al. Predicting poor outcomes for back pain seen in primary care using patients' own criteria. *Spine* 1996;21:2900–7.
- Blier F, Abbott FV. Putative mechanisms of action of antidepressant drugs in affective and anxiety disorders and pain. *J Psychiatry Neurosci* 2001;26:37–43.
- Lynch ME. Antidepressants as analgesics: a review of randomized controlled trials. *J Psychiatry Neurosci* 2001;26:30–6.
- Watson JC, Gordon LB, Stermac L, et al. Comparing the effectiveness of process experiential with cognitive-behavioral psychotherapy in the treatment of depression. *J Consult Clin Psychol* 2003;71:773–81.
- Lin EHB, Katon W, Von Korff M, et al. Effect of improving depression care on pain and functional outcomes among older adults with arthritis: A randomized controlled trial. *JAMA* 2003;290:2428–34.
- Parker JC, Smarr KL, Slaughter JR, et al. Management of depression in rheumatoid arthritis: A combined pharmacologic and cognitive-behavioral approach. *Arthritis Rheum* 2003;49:766–77.
- Rost K. Depression and pain: The road to evidence based care. *Arch Intern Med* 2003;163:2415–6.