



Published in final edited form as:

*Spine J.* 2013 November ; 13(11): 1421–1433. doi:10.1016/j.spinee.2013.05.047.

## One-year Outcomes of Surgical versus Non-surgical Treatments for Discogenic Back Pain: A Community-based Prospective Cohort Study

Sohail K. Mirza, MD, MPH<sup>1</sup>, Richard A. Deyo, MD, MPH<sup>2</sup>, Patrick J. Heagerty, PhD<sup>3</sup>, Judith A. Turner, PhD<sup>4</sup>, Brook I. Martin, PhD<sup>1</sup>, and Bryan A. Comstock, MS<sup>3</sup>

<sup>1</sup>Dartmouth Medical School, Lebanon, N.H. 03756

<sup>2</sup> Department of Family Medicine, Medicine, Public Health and Preventive Medicine, and the Center for Research in Occupational and Environmental Toxicology, Oregon Health and Science University. 3181 SW Sam Jackson Park Rd. Portland, OR 97239-3098

<sup>3</sup> Department of Biostatistics, University of Washington, Box 357232, 1959 NE Pacific Street, Seattle, WA 98195

<sup>4</sup> Department of Psychiatry and Behavioral Sciences, University of Washington, Box 356560, 1959 NE Pacific Street, Seattle, WA 98195

### Introduction

Back pain associated with disc degeneration may be the most controversial subject in spine care, and perhaps the one most in need of further clinical research. Because chronic back pain is so disabling and so common<sup>1,2</sup>, a large population of vulnerable patients is yearning for any promise of relief. These patients are attracted to an expanding range of costly diagnostic and therapeutic interventions.<sup>2,3</sup> They may be unaware of the controversy surrounding many aspects of a diagnosis of “discogenic back pain.”<sup>4</sup> This diagnosis does not

**Correspondence:** Sohail K. Mirza, M.D., M.P.H. Chair, Department of Orthopaedics Dartmouth-Hitchcock Medical Center Lebanon, NH 03756 Phone: 603 653-6090 Fax: 603 653-3581 mirza@dartmouth.edu.

Sohail K. Mirza mirza@Dartmouth.edu

Richard A. Deyo deyor@ohsu.edu

Patrick J. Heagerty heagerty@u.washington.edu

Judith A. Turner jturner@u.washington.edu

Brook I. Martin Brook.I.Martin@Dartmouth.edu

Bryan A. Comstock bac4@u.washington.edu

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Access to Patient Level Data

The NIH/NIAMS Multidisciplinary Clinical Research Center (MCRC) at Dartmouth will provide access to a de-identified limited data set containing patient level data for this study. Any investigator is eligible for access to a flat file containing data on all 495 patients enrolled in the study and the accompanying data dictionary. The investigator is required to submit an analysis protocol defining the objective, aims, and analysis plan. The MCRC methodology core will review the plan for clarity and completeness. After a plan is approved, the investigator will then be granted access to download the data and the accompanying data dictionary from the MCRC website. The original study plan, comments from the MCRC methodology review, and the revised study plan will be posted for open access on the MCRC website. Any manuscripts, presentations, or other dissemination resulting from the analysis will be required to reference this study plan on the MCRC website.

have well-established criteria. It generally refers to back pain in patients without radicular symptoms and without structural abnormalities other than lumbar disc degeneration. Professional societies provide some guidelines on this condition<sup>5-8</sup>, but no consensus exists on whether or how to follow the recommendations.<sup>9-12</sup> It remains unclear, for example, whether there is a localized peripheral generator of back pain<sup>13</sup> (i.e., the intervertebral disc), whether imaging studies such as magnetic resonance imaging (MRI)<sup>14</sup> or diagnostic tests such as discography<sup>15</sup> can distinguish painful from non-painful discs, or whether such a pain source can be eliminated by non-surgical treatment, excision, fusion, or artificial disc replacement.

Lack of consensus regarding the efficacy of lumbar spinal fusion for discogenic back pain is particularly troubling because four European randomized trials have compared fusion to non-surgical care.<sup>16-19</sup> Fusion showed a small benefit for back disability when compared to non-standardized non-surgical care, but roughly similar benefit when compared to intensive rehabilitation incorporating cognitive-behavior therapy.<sup>20</sup> The Food and Drug Administration Investigational Device Exemption (FDA-IDE) artificial disc approval studies showed disc replacement to have less than 60% success rate for a composite outcome, and even lower success for the comparator, lumbar fusion.<sup>21-23</sup> In the context of rising utilization and costs of lumbar fusion,<sup>24,25</sup> these results have invited scrutiny from payers.<sup>26,27</sup> Independent evidence reviews commissioned by the Centers for Medicare and Medicaid Services Coverage and Advisory Committee (MCAC)<sup>28</sup> and the Washington Healthcare Technology Assessment Program<sup>29</sup> concluded that lumbar fusion for degenerative disc disease lacked sufficient evidence of efficacy and safety to justify unconditional coverage, and private payers have reached similar conclusions for artificial disc replacement.<sup>30</sup> In contrast, experienced surgeons claim “properly selected patients” have successful outcomes with surgery, without specifying the selection criteria.<sup>31</sup>

The randomized trials were performed in countries with nationalized healthcare. We considered the possibility that patients in United States with “discogenic back pain” who receive surgery in a community-based practice setting may differ in important, measurable characteristics from those who do not receive surgery; they may also have different outcomes. They may have different expectations, different access to services, and different social support options. Also, in a controlled trial, the strict eligibility criteria and standardized interventions may not reflect real-world practice. An observational study to describe the demographics, baseline features, and treatment utilization among US patients with discogenic back pain would add new and complementary information to these trials. We designed an observational prospective cohort study to address these hypotheses.<sup>32</sup> Our goals were to select patients presenting for initial surgical consultation for axial back pain associated with disc degeneration, select those patients the treating surgeon identified as having “discogenic back pain”, identify baseline characteristics associated with receiving surgery, and compare outcomes of surgical versus non-surgical treatment. In contrast to a controlled trial, we did not interfere with the diagnosis and care process; we simply did our best to record what was done and how the patients’ pain and function changed.

## Methods

### Study Design

We conducted a prospective cohort study of patients with axial back pain seeking surgical consultation. The detailed study protocol has been previously published.<sup>32</sup> To obtain a representative sample of community practice, we enrolled patients from five sites: a county hospital, an academic medical center, and three community hospitals. Orthopedic surgeons (n=12) and neurosurgeons (n=4) participated. The study protocol was approved by the University of Washington institutional review board (IRB) and the IRBs of participating community hospitals, and all study participants provided written informed consent. Study participants were followed for 12 months following enrollment.

### Patient Selection

Our goal was to identify patients with “discogenic back pain” consulting a surgeon for the first time to discuss surgical treatment options. Because no established criteria exist for this diagnosis, we relied on the judgment of the treating surgeon regarding interpretation of the patient's clinical presentation and imaging studies. We required that patients have low back pain as the primary symptom and an MRI scan confirming disc degeneration at only one or two lumbar discs. Enrollment criteria did not restrict the MRI to be within any specific time interval. The diagnosis of discogenic back pain was established by the surgeon. Investigators reviewed the radiologist's report to confirm that no specific structural abnormalities were reported by the radiologist and degeneration was limited to 1 or 2 levels. Because discography remains controversial,<sup>33</sup> we did not require discography as an enrollment criterion. We required symptom duration of at least 6 months, as surgeons rarely consider surgery for discogenic pain of shorter duration. We did not ask whether the pain varied during the duration interval (e.g. pain every day, or most days).

Research coordinators screened records of all patients presenting with back pain at the participating sites. Coordinators used a pre-specified list of exclusion criteria. Patients could be excluded at screening, baseline interview, or subsequent confirmation of enrollment criteria and consent. We excluded patients with neurological deficit or predominantly nerve root symptoms, motor deficits, abnormal electrodiagnostic studies (if performed), structural spine deformity such as stenosis or spondylolisthesis (> Grade I), inflammatory disease, spinal malignancy, instability on radiographs (if performed), pregnancy, other specific causes for back pain, or severe comorbidity that would contraindicate surgery. We also excluded patients older than 65 years because of the high prevalence of multilevel disc degeneration and spinal stenosis in this age group.

### Definition of Treatment Groups

In this observational study, we did not specify the treatments patients received. At each assessment, we asked patients about treatments they had received. We also recorded treatments listed in their medical records. Patients who underwent surgery within 6 calendar months after study enrollment date were designated as “surgical”; those who did not were designated as “non-surgical”, even if they had surgery later in the study period. Patients could undergo multiple treatments concurrently during the study period. Any additional or

concurrent non-operative treatments received by surgical patients during the initial six months after enrollment were considered co-interventions.

### Outcome Measures

Our primary outcome was the modified Roland-Morris Back Disability score as measured by the number of “yes” responses to 23 statements describing activity limitations related to back pain.<sup>34</sup> Higher scores indicate greater disability. A 5-point or 30% reduction from baseline score<sup>35,36</sup> is considered the minimal clinically important change in this score.

Secondary outcome measures included patient current rating of overall pain severity on a numerical scale of 0 = “no pain” to 10 = “worst possible pain”, back and leg pain bothersomeness measures,<sup>37-39</sup> the Physical Function scale of the Short Form 36 version 2 (SF-36v2) questionnaire,<sup>40</sup> use of medications for pain, and work status.

We also examined success rate using a composite definition of success: 30% improvement from baseline in the Roland score, 30% improvement from baseline in current pain rating, no opioid medication use within the past three months, and working (for patients for whom work was relevant, i.e. not retired, working in the home, or receiving disability compensation prior to surgery).

Outcomes were assessed in person at baseline and by telephone interviews or mailed questionnaires at 3, 6, 9, and 12 months after enrollment. Interviewers were blinded to the subject’s treatment group designation. Study participants were considered lost for a particular follow-up after a minimum of 12 dispersed unsuccessful telephone interview attempts and no response to three sequential mailings during the follow-up time window.

### Baseline measures

At baseline, we assessed a variety of patient characteristics in order to describe the sample. These included patient sociodemographic characteristics, history of symptoms and treatments, work status, work disability compensation status, and litigation.<sup>32</sup> We recorded medical comorbidity using a questionnaire based on the Charlson comorbidity index<sup>41</sup> and asked about smoking, alcohol and drug use<sup>42</sup>. Psychological measures included the Symptom CheckList 90 Depression and Somatization scales,<sup>43</sup> the SF-36, version 2, Mental Health scale, and the Pain Catastrophizing Scale.<sup>44</sup> Psychological measures were also administered at follow-up during the first year of the study, but patient complaints of questionnaire burden required us to modify our protocol and collect these only at baseline.

### Assessment of Therapeutic Safety

Because patients were recruited from multiple practices in this community-based study, we chose to evaluate therapeutic safety primarily through information obtained by patient interviews. We selected three adverse outcomes common to both surgical and non-surgical treatment groups: emergency department (ED) visits, hospitalizations, repeat surgery in the surgical group, any surgery after the treatment-group designation period (first six-months following enrollment) for the non-surgical group. These measures were ascertained uniformly for both groups through patient interviews. We reviewed operation reports and

hospital discharge summaries of the surgical patients. We also performed detailed safety surveillance of surgical patients at two hospitals,<sup>45</sup> but we did not have sufficient study personnel for this type of direct hospitalization surveillance of all study participants.

### Statistical Analysis

Our sample size calculations indicated we needed approximately 95 patients in each treatment arm to detect a 3.0 point difference in the Roland Score (primary outcome), based on a standard deviation of 7.37 from the Maine Lumbar Spine Study,<sup>46,47</sup> two-sided alpha = 0.05, and power = 0.80.<sup>32</sup> Since observational analyses have a strong potential for meaningful biases, we present confidence levels and sensitivity analyses for observed differences for clinically important baseline features and outcome measures.

To compare the primary outcome (1-year modified Roland score) in the two treatment groups, we used a linear mixed-effects regression model with treatment status entered as an independent variable. The model used available information and multiple imputation to account for missing data and patients lost to follow-up. Multiple imputation of missing data points is germane to the repeated outcome measures of a linear mixed-effects regression model. Imputed data points are derived from the baseline covariates in the model, and the correlation of each patients recorded outcomes over time. In essence, we used baseline characteristics and any available outcome data to estimate the most likely missing data points, based on data from similar patients. We included random effects to account for correlated measures collected on the same individuals over time with unstructured covariance. Using bivariate and multivariate logistic regression models, we examined the associations between baseline characteristics and having surgery to assess the risk of confounding in this non-randomized study. We adjusted the linear mixed effects model for potentially confounding baseline factors that were significantly associated with receipt of surgery.

The inclusion of covariates for the linear mixed model using  $p < 0.05$  for prediction of surgery may be too strong and exclude some other factors that are indeed associated with surgery, for which control could impact estimates. Therefore, as a sensitivity measure for the primary outcome, we also performed an alternative analysis using a stepwise selection logistic regression method that included baseline variables for outcomes (Roland, SF-36 (8 domains), catastrophizing, depression, somatization, helplessness and rumination, back and leg pain), patient characteristics (gender, education, race, age, work status, back pain duration, body mass index, comorbidity, smoking, alcohol, marital status, disability, and lawyer help), and resources (days cut down on activity, bed days, missed work, prescriptions). We set the probability to remove at  $>0.20$ . The following variables remained in the model: SF-36 vitality, social function, role physical, and general health domains, leg weakness, work status, cut down on activities, somatization, catastrophizing, helplessness, recruitment site, symptom duration, and previous surgery.

The number of interventions received during the treatment period was compared between groups using a trend test (ptrend in Stata). Proportions for the categorical outcome measures (safety and the composite definition of success) were compared using logistic regression with adjustment for covariates. All analyses were performed using STATA version 11.0

(StataCorp LP, College Station, TX). All tests were two-sided and p-values < 0.05 were considered to be significant.

## Results

### Study participants

Figure 1 shows the study flow. We screened 7,344 patients who had back pain and were referred to the study. Of these, 495 met the study inclusion criteria and agreed to participate. Reasons for exclusions are listed in Table 1. The most frequent exclusions were for age greater than 65 years (922, 13.5%), presence of radiculopathy (848, 12.4%), problem not related to lumbar spine (827, 12.1%), and prior fusion (783, 11.4%). Our sample size target was 95 in each treatment group, but after nearly 5 years of enrollment, only 86 (17%) of 495 study participants received surgery within 6 months of the enrollment date. Although more patients in the surgical group would have been optimal, we stopped further enrollment because of limits in our project funding.

Twelve-month post-enrollment interviews were completed in 70 (81%) surgical patients and 336 (82%) non-surgical patients (Figure 1). Patients lost at the 12-month post-enrollment interview were significantly more likely to have baseline characteristics generally associated with worse outcomes: more prior surgery, more severe pain, greater physical disability, and more worker's compensation claims.

### Baseline Comparisons of Surgical and Non-surgical Patients

Surgical and non-surgical patients were similar in most of the characteristics we measured at baseline, including measures of psychological distress (Table 2). However, patient with prior lumbar decompression surgery, greater baseline back and leg pain bothersomeness, and greater back-related physical disability were more likely to receive surgery in the next six months (Table 2). In the multivariate model, after adjusting for other important baseline characteristics, prior surgery, greater back-related physical disability, and being seen at a private practice site (as compared with an academically-affiliated hospital) were associated with receiving surgery. We also found a trend towards a significantly lower chance of receiving surgery for patients who were smokers, and a trend towards receiving surgery for patients with greater leg pain bothersomeness, controlling for other factors (not shown).

There was a wide range of scores on each baseline measure. On average, study participants scored almost one standard deviation below (i.e., worse than) the general population mean on the SF-36 mental health scale, showed moderate levels of depressive symptoms, and reported multiple nonspecific physical symptoms. Mean pain-related catastrophizing scores were similar to those reported among patients seen in outpatient pain clinics.<sup>48,49</sup>

### Nature of the Surgical Treatments and Co-interventions

Surgical treatment varied, consisting of instrumented fusion in 68 patients (79% of the surgical patients), artificial disc replacement in 10 (12%), and laminectomy or discectomy in 8 (9%). Enrollment criteria were confirmed in these patients; we cannot explain the rationale for decompression surgery. Surgical patients also received multiple non-surgical co-



interventions during the first 6 months following enrollment (Table 3). The number of non-operative treatments was greater in the surgical group than in the non-surgical group within the first six months ( $p=0.01$ , Table 3). Surgery was performed at a mean of 2.4 months after enrollment.

### Nature of Non-surgical Treatments

Non-surgical treatments received by patients during the first 6 months of the study were similar to the treatments patients reported receiving prior to enrollment, although rates of use were lower (Table 3). A substantial minority of patients (13%) reported receiving no treatment during the first six months following enrollment, the treatment designation period for the study.

### Outcomes of Surgical and Non-surgical Treatment

The primary outcome, back-specific disability, showed advantage for surgery (Figure 2). Based on linear mixed models adjusting for baseline measures associated with receipt of surgery and for loss to follow-up, patients who received surgery in the first six months of the study, on average, had Roland scores that were 5.4 points (95% CI 3.9 to 6.9,  $p < 0.001$ ) lower than those of patients in the non-surgical group at one year (6 to 12 months following surgery; mean 9.6 months after surgery).

Alternative analysis using a stepwise model suggested a slightly greater improvement in the surgical group than the original model, but this analysis was not as parsimonious, and lost some subjects due to missing baseline variables. The Roland score for the surgical group at 12 months was 6.07 points lower (95% CI 4.4 - 7.8) compared to the non-surgical group ( $p < 0.001$ ), overlapping with the estimate from the original model. However, the stepwise model included many more terms and only 399 subjects due to missingness; the original model included 495 subjects and fewer parameters. The statistical tests between surgery and non-surgery were not otherwise different, and the conclusions remain the same.

Secondary outcomes also showed advantage for surgery, including overall pain intensity rating (Figure 2), composite measures of success (Table 4), and other Physical and Mental Health measures (Table 5). Patients showed variable improvement on both the overall pain rating and Roland back disability scales, but the surgical group had greater improvement than the non-surgical group at all potential cut-off thresholds for defining success (Figure 3). Using a composite measure of success defined as 30% improvement from baseline in the Roland score, 30% improvement from baseline in pain rating, return to work for eligible workers, and no opioid pain medication use, the 1-year post-enrollment success rate was 33% in the surgery group and 15% in the non-surgical group ( $p < 0.001$ ; Table 4). Some patients did not improve: approximately 25% of non-surgical and 15% of surgical patients reported worse function and increased pain at 12-months following enrollment compared to baseline (Figure 3).

### Patient-Reported Measures of Safety

Between 6 and 12 months after enrollment, ED visits occurred with similar frequency in the two treatment groups: 5/76 (7%) of surgical patients and 41/366 (11%) non-surgical patients

( $p=0.23$ ). Overnight hospitalizations also occurred with similar frequency in both groups: 1/76 (1) surgical and 13/366 (4%) non-surgical ( $p=0.31$ ). Repeat surgery occurred in 8/76 (11%) of surgical patients. Also, 22/366 (6%) of patients in the non-surgical group received surgery between 6 -12 months after enrollment.

## Discussion

This community-based comparative effectiveness study showed only fair outcomes for both surgical and non-surgical treatment of discogenic back pain. Patients with chronic back pain who seek surgical consultation and are found to have “discogenic back pain” presented, on average, with moderate levels of pain, physical disability, and psychological distress. When assessed 12 months after enrollment, patients who had surgery combined with non-surgical co-interventions showed modest but significantly greater improvement in self-reported disability, back pain, generic physical function, and composite success measures compared to patients who had continuation of unstructured non-surgical care. Surgical patients also concurrently received more intensive co-treatments than the non-surgical group. Although surgery combined with various additional non-surgical treatments showed advantage over non-surgical treatment alone, only one-third of surgical patients attained a successful result defined by stringent criteria of clinically-important improvement in pain and function, no opioid medication use, and return to work for eligible workers. The rate of activity restriction and opioid use was significantly greater in the surgical group along with greater use of corsets and bed rest. Surgery did not reduce the frequency of emergency room visits or overnight hospitalizations, and 11% of patients had repeat surgery within the first year post-operatively.

Patients, on average, showed minimal improvement after 12-months of continued non-surgical care as currently provided in the US. Non-surgical treatments patients received varied widely, were utilized in an unstructured manner, and mostly did not comply with conservative care guidelines. For example, only 5% received cognitive behavior therapy. Outcomes in the non-surgical group may have improved more if treatment had adhered to recommendations from clinical practice guidelines.<sup>6</sup>

The poor prognosis we observed for discogenic back pain has policy implications. The non-surgical patients in our study provide somewhat of a natural history of discogenic back pain since they received minimal new therapy after enrollment. In contrast to the commonly held view that non-specific back pain has a benign course, the prognosis for those who sought surgical consultation had marked pain and functional limitations at enrollment and remained essentially unchanged during 12-months of surveillance. The fundamental concepts underlying the mechanism of pain in these patients, the treatment options offered to them, and policies governing these treatments need re-evaluation. New diagnostic and surgical technologies are readily available in community practice, but comprehensive rehabilitation and cognitive behavior therapy are difficult to find and frequently not covered by insurance programs.

We conducted an observational study to obtain a “real-world” or pragmatic view of community-based practice and outcomes for discogenic back pain. We acknowledge that



observational studies comparing treatment effectiveness have many limitations; results should be interpreted with caution. Despite statistical adjustments, unmeasured confounding factors persist and bias observed associations. The population chosen for this study was likely biased in favor of surgery since it had already made the decision to seek a surgical consultation. Only a small fraction of patients with back pain seeking surgical opinion were judged by the treating surgeon to have discogenic back pain (i.e. chronic back pain, disc degeneration at only one or two lumbar levels, and no other focal abnormalities).

In contrast to the standardization often imposed in randomized trials, this study shows that under natural conditions, patients with chronic back pain often mix multiple treatment interventions concurrently. Surgical patients simultaneously received multiple non-surgical co-interventions. In fact, patients who underwent surgery received more non-operative treatments than did patients in the non-surgical group. We enrolled patients after the Fritzell<sup>50</sup> and Brox<sup>51</sup> studies were published. During orientation of surgeons participating in our study, we reviewed data showing intensive rehabilitation incorporating cognitive behavior therapy was just as effective as surgery.<sup>52</sup> Despite availability of this knowledge, the non-surgical care received by patients in our study remained haphazard.

Our study reports early, short-term results. Follow-up for surgical patients averaged 9.6 months post-operatively. Lumbar arthrodesis procedures can require that duration for healing, placebo effects may be particularly strong in the early post-operative period,<sup>53</sup> and the advantage for surgery may diminish with time. Surgical advantage was greatest at 1 year and diminished by 2 years in a Swedish randomized trial comparing lumbar fusion to non-operative care;<sup>19</sup> outcomes for surgical and non-surgical groups were similar at 5-year follow-up.<sup>54</sup> A small Japanese trial also showed early advantage for surgery.<sup>55</sup>

Although it is widely accepted that back pain in some patients may be caused by lumbar disc degeneration, the diagnosis of discogenic back pain lacks a firm biological basis and clear clinical description. It is uncertain whether the structural and physiological intervertebral disc changes associated with aging alone can be distinctly separated from changes that cause low back pain. Because disc degeneration is almost ubiquitous beyond the age of 50, and because back pain is very common, this ambiguity has important clinical implications. Some physicians believe that individual discs can be identified as sources of pain in individual patients, and infer that surgery to immobilize or replace the disc will help eliminate back pain.<sup>56</sup> Others believe it is nearly impossible to identify specific discs as a cause of pain in individual patients.<sup>56</sup> A Combined Task Force of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology for Nomenclature & Classification of Lumbar Disc Pathology acknowledged the difficulty of distinguishing disc degeneration from normal aging.<sup>57</sup>

Efficacy cannot be proven by non-randomized studies such as ours.<sup>58</sup> We address group differences by reporting multiple comparisons of baseline measures, interventions, and outcomes. However, confounding by unmeasured factors may account for the associations observed between surgery and outcomes. Surgical patients in our study also received more intensive non-operative co-interventions compared to the non-surgical cohort, none of whom received structured, state-of-the-art rehabilitation, and some received no treatment.

However, if patient preferences, the treating surgeon's patient selection biases, co-interventions, and other unmeasured factors associated with receiving surgery are considered as a bundled package (i.e. “use effectiveness”<sup>59</sup>), surgical patients had better - but still poor - outcomes compared to continued unstructured non-surgical care.

Our study shows that patients in whom a spine surgeon labeled as having discogenic back pain have severe functional limitations at baseline, have multiple comorbidities, and receive multiple concurrent treatments in course of usual care. In contrast to the general perception that patients with non-specific back pain improve with minimal treatment, we found these patients continue to have severe pain and functional limitations. The degree of improvement we observed in surgical patients was marginal even with this imbalance of treatment intensity. These findings are relevant to guiding policy and practice for this patient population in the United States. Following even the most basic conservative care guidelines may yield large improvements.

## Acknowledgements

This work was supported by grants from the National Institutes of Health/National Institute of Arthritis, Musculoskeletal, and Skin Disorders 5K23AR48979, 5P60-AR48093, and RC1AG036268. It was also supported in part by the Spine End-Results Research Fund at the University of Washington established through a gift from Synthes Spine (Paoli, PA), a manufacturer of spinal surgery implants.

We are grateful to the following surgeons and their office staff for referring patients to the study and providing valuable feedback:

- Bone and Joint Center, University of Washington: Dheera Ananthakrishnan, Rick Bransford, Jens Chapman, Todd Jarosz, Ted Wagner
- Evergreen Hospital: Reginald Knight
- Harborview Medical Center: Carlo Bellabarba, Jens Chapman, Andrew Dailey, Alex West
- Providence Medical Center: David Hanscom, Paul Schwaegler
- Swedish Medical Center: Jeff Garr, Jay Williams
- University of Washington Medical Center: Christopher Shaffrey, Trent Tredway

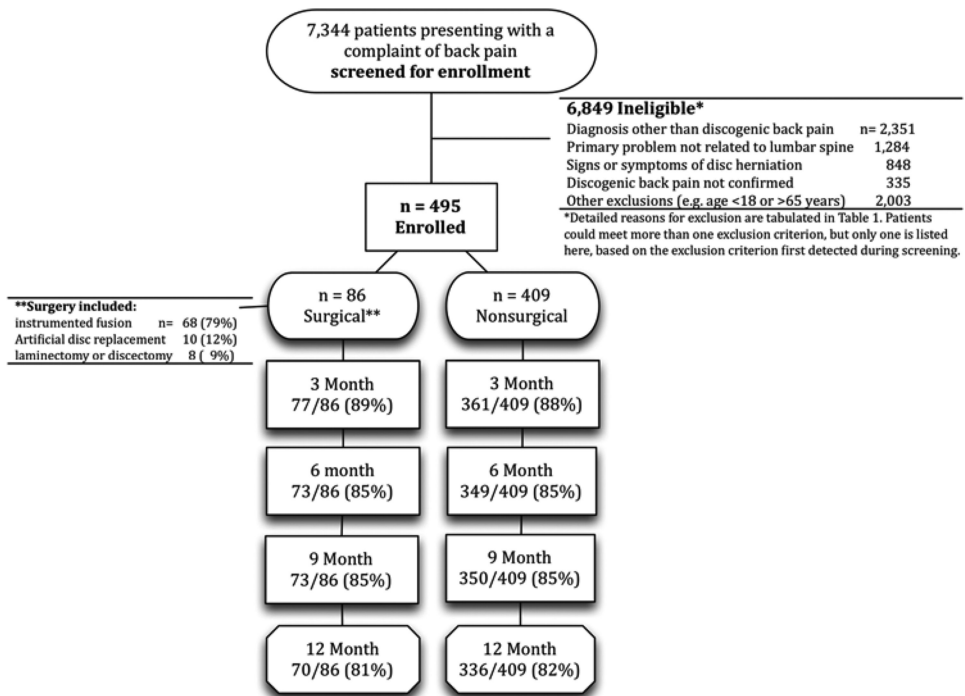
## References

1. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine*. 2006; 31:2724–7. [PubMed: 17077742]
2. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA : the journal of the American Medical Association*. 2008; 299:656–64. [PubMed: 18270354]
3. Ivanova JI, Birnbaum HG, Schiller M, Kantor E, Johnstone BM, Swindle RW. Real-world practice patterns, health-care utilization, and costs in patients with low back pain: the long road to guideline-concordant care. *The spine journal : official journal of the North American Spine Society*. 11:622–32. [PubMed: 21601533]
4. Groopman J. A knife in the back: is surgery the best approach to chronic back pain? *The New Yorker*. 2002
5. Chou R, Loeser JD, Owens DK, et al. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. *Spine*. 2009; 34:1066–77. [PubMed: 19363457]
6. Chou R, Atlas SJ, Stanos SP, Rosenquist RW. Nonsurgical interventional therapies for low back pain: a review of the evidence for an American Pain Society clinical practice guideline. *Spine*. 2009; 34:1078–93. [PubMed: 19363456]

7. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD. Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. *Spine*. 2009; 34:1094–109. [PubMed: 19363455]
8. Resnick DK. Evidence-based guidelines for the performance of lumbar fusion. *Clin Neurosurg*. 2006; 53:279–84. [PubMed: 17380763]
9. Chou R, Chou R, Loeser JD, Owens DK, et al. Interventional therapies, surgery and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. *Spine*. 2009; 34:1066–77. *Spine*;36:590. [PubMed: 19363457]
10. Datta S, Manchikanti L, Chou R, Loeser JD, Owens DK, et al. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. *Spine*. 2009; 34:1066–77. *Spine*;35:1826; author reply -7. [PubMed: 19363457]
11. Manchikanti L, Datta S, Derby R, Wolfer LR, Benyamin RM, Hirsch JA. A critical review of the American Pain Society clinical practice guidelines for interventional techniques: part 1. Diagnostic interventions. *Pain Physician*. 13:E141–74. [PubMed: 20495596]
12. Furlan AD, Pennick V. Limitations of guidelines for low back pain therapy. *Nat Rev Rheumatol*. 2009; 5:473–4. [PubMed: 19710667]
13. Giesecke T, Gracely RH, Grant MA, et al. Evidence of augmented central pain processing in idiopathic chronic low back pain. *Arthritis Rheum*. 2004; 50:613–23. [PubMed: 14872506]
14. Carragee E, Alamin T, Cheng I, Franklin T, van den Haak E, Hurwitz E. Are first-time episodes of serious LBP associated with new MRI findings? *Spine J*. 2006; 6:624–35. [PubMed: 17088193]
15. Carragee EJ, Alamin TF, Carragee JM. Low-pressure positive Discography in subjects asymptomatic of significant low back pain illness. *Spine*. 2006; 31:505–9. [PubMed: 16508542]
16. Brox JI, Reikeras O, Nygaard O, et al. Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. *Pain*. 2006; 122:145–55. [PubMed: 16545523]
17. Brox JI, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine*. 2003; 28:1913–21. [PubMed: 12973134]
18. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *Bmj*. 2005; 330:1233. [PubMed: 15911537]
19. Fritzell P, Hagg O, Wessberg P, Nordwall A. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine*. 2001; 26:2521–32. discussion 32-4. [PubMed: 11725230]
20. Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine*. 2007; 32:816–23. [PubMed: 17414918]
21. Blumenthal S, McAfee PC, Guyer RD, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. *Spine*. 2005; 30:1565–75. discussion E387-91. [PubMed: 16025024]
22. McAfee PC, Cunningham B, Holsapple G, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. *Spine*. 2005; 30:1576–83. discussion E388-90. [PubMed: 16025025]
23. Zigler J, Delamarter R, Spivak JM, et al. Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine*. 2007; 32:1155–62. discussion 63. [PubMed: 17495770]

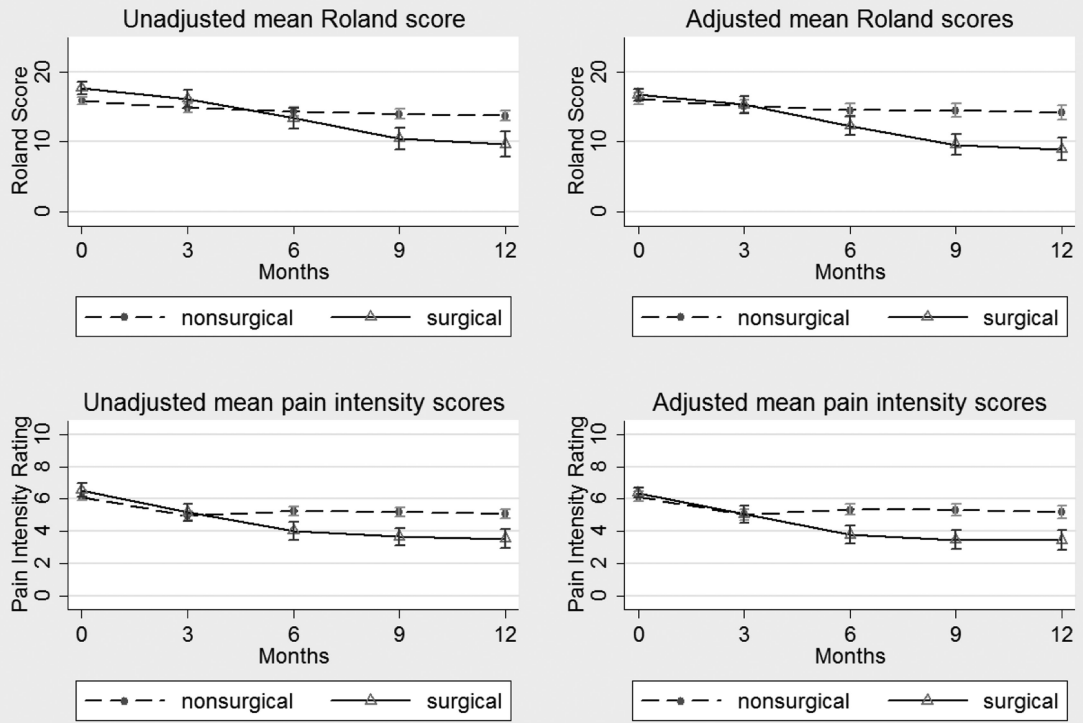
24. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. *Spine*. 2005; 30:1441–5. discussion 6-7. [PubMed: 15959375]
25. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery - the case for restraint. *The New England journal of medicine*. 2004; 350:722–6. [PubMed: 14960750]
26. CMC/HHS. Medicare Program; Meeting of the Medicare Coverage Advisory Committee-- November 30, 2006. Vol. 71. Federal Register; 2006. Accessed at <http://a257gakamaitech.net/7/257/2422/01jan20061800/edocketaccessgpo.gov/2006/E6-13938.htm> on Nov 27, 2007
27. Health Technology Clinical Committee Meeting Schedule November 16, 2007 Lumbar Fusion and Discography. accessed at <http://www.hta.hca.wa.gov/committee/index.shtml> on Nov 27, 2007. Washington State Healthcare Authority, 2007. (Accessed at
28. Spinal Fusion for Treatment of Degenerative Disease Affecting the Lumbar Spine. Agency for Healthcare Research and Quality; 2006. accessed at <http://www.cms.hhs.gov/determinationprocess/downloads/id41ta.pdf> on Nov 27, 2007(Accessed at <http://www.cms.hhs.gov/determinationprocess/downloads/id41ta.pdf>.)
29. Spinal Fusion and Discography For Chronic Low Back Pain and Uncomplicated Lumbar Degenerative Disc Disease. Washington Health Technology Assessment; 2007. accessed at [http://www.hta.hca.wa.gov/docs/spinal\\_fusion\\_discography\\_final\\_101907.pdf](http://www.hta.hca.wa.gov/docs/spinal_fusion_discography_final_101907.pdf) on Nov 27, 2007(Accessed at
30. BCBS. Artificial lumbar disc replacement. Technol Eval Cent Asses Program Exec Summ. 2007; 22:1–3.
31. Errico TJ, Gatchel RJ, Schofferman J, et al. A fair and balanced view of spine fusion surgery. *Spine J*. 2004; 4:S129–38. [PubMed: 15374548]
32. Deyo RA, Mirza SK, Heagerty PJ, Turner JA, Martin BI. A prospective cohort study of surgical treatment for back pain with degenerated discs; study protocol. *BMC musculoskeletal disorders*. 2005; 6:24. [PubMed: 15913458]
33. Wichman HJ. Discography: over 50 years of controversy. *Wmj*. 2007; 106:27–9. [PubMed: 17393754]
34. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995; 20:1899–908. discussion 909. [PubMed: 8560339]
35. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*. 2008; 33:90–4. [PubMed: 18165753]
36. Jordan K, Dunn KM, Lewis M, Croft P. A minimal clinically important difference was derived for the Roland-Morris Disability Questionnaire for low back pain. *J Clin Epidemiol*. 2006; 59:45–52. [PubMed: 16360560]
37. Keller RB, Atlas SJ, Singer DE, et al. The Maine Lumbar Spine Study, Part I. Background and concepts. *Spine*. 1996; 21:1769–76. [PubMed: 8855461]
38. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine*. 1996; 21:1777–86. [PubMed: 8855462]
39. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. *Spine*. 1996; 21:1787–94. discussion 94-5. [PubMed: 8855463]
40. Ware JE Jr. SF-36 health survey update. *Spine*. 2000; 25:3130–9. [PubMed: 11124729]
41. Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical record review? *Med Care*. 1996; 34:73–84. [PubMed: 8551813]
42. Brown RL, Leonard T, Saunders LA, Pappasoulis O. A two-item conjoint screen for alcohol and other drug problems. *J Am Board Fam Pract*. 2001; 14:95–106. [PubMed: 11314930]
43. Derogatis LR, Rickels K, Rock AF. The SCL-90 and the MMPI: a step in the validation of a new self-report scale. *Br J Psychiatry*. 1976; 128:280–9. [PubMed: 1252693]
44. Sullivan MJL, Bishop SR, Scott R, Pivik J. The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment*. 1995; 7:524–32.
45. Mirza SK, Deyo RA, Heagerty PJ, Turner JA, Lee LA, Goodkin R. Towards standardized measurement of adverse events in spine surgery: conceptual model and pilot evaluation. *BMC musculoskeletal disorders*. 2006; 7:53. [PubMed: 16787537]

46. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine*. 1996; 21:1777–86. [PubMed: 8855462]
47. Atlas SJ, Deyo RA, van den Ancker M, Singer DE, Keller RB, Patrick DL. The Maine-Seattle back questionnaire: a 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis: results of a derivation and validation cohort analysis. *Spine*. 2003; 28:1869–76. [PubMed: 12923478]
48. Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The Pain Catastrophizing Scale: further psychometric evaluation with adult samples. *J Behav Med*. 2000; 23:351–65. [PubMed: 10984864]
49. Sullivan MJ, Stanish W, Waite H, Sullivan M, Tripp DA. Catastrophizing, pain, and disability in patients with soft-tissue injuries. *Pain*. 1998; 77:253–60. [PubMed: 9808350]
50. Fritzell P, Hagg O, Wessberg P, Nordwall A. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine*. 2001; 26:2521–32. discussion 32–4. [PubMed: 11725230]
51. Brox JI, Sorensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine*. 2003; 28:1913–21. [PubMed: 12973134]
52. Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine*. 2007; 32:816–23. [PubMed: 17414918]
53. Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. *Jama*. 2006; 296:2483–5. [PubMed: 17119146]
54. Fritzell P. Annual meeting of the International Society for Study of the Lumbar Spine, Porto, Portugal. 2004 2004.
55. Ohtori S, Koshi T, Yamashita M, et al. Surgical versus nonsurgical treatment of selected patients with discogenic low back pain: a small-sized randomized trial. *Spine*. 36:347–54. [PubMed: 20838371]
56. Nachemson A, Zdeblick TA, O'Brien JP. Lumbar disc disease with discogenic pain. What surgical treatment is most effective? *Spine*. 1996; 21:1835–8. [PubMed: 8855471]
57. Fardon DF, Milette PC. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. *Spine*. 2001; 26:E93–E113. [PubMed: 11242399]
58. Weiss, NS. *Clinical Epidemiology: the study of the outcome of illness*. Second Edition ed.. Oxford University Press; New York: 1996.
59. Meier P. Comment on “Compliance as an explanatory variable in clinical trials” by Efron and Feldman. *J Am Stat Assoc*. 1991; 86:19–22.



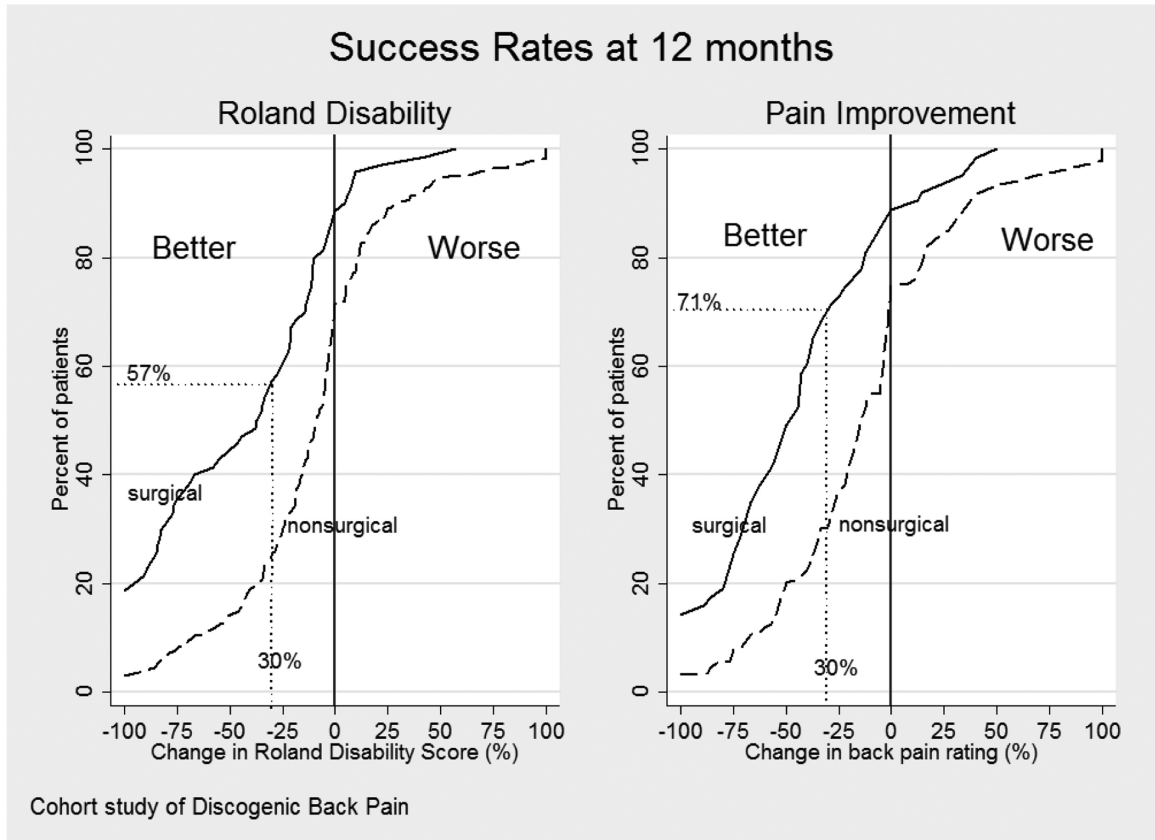
**Figure 1.**  
Flow diagram of patient screening, eligibility, enrollment, and follow-up.





error bars represent 95% confidence intervals

**Figure 2.** Adjusted and unadjusted Roland disability (primary outcome) and back pain intensity at each follow-up time point measured from enrollment date.



**Figure 3.** Cumulative proportion of responders for the 12-month change in the Roland disability score (primary outcome) and overall pain intensity rating over the entire range of possible cut-off points for defining success. The graphs allow comparison of the treatment groups at any response level that a surgeon or patient may consider important. For example, if 30% or more improvement from baseline is considered the criterion for success, 57% of surgical patients achieved it on the Roland Scale compared to 25% of non-surgical patients. For overall pain (0 to 10 numerical rating scale), 71% of the surgical patients and 30% of the non-surgical patients achieved 30% or more improvement compared to baseline scores.

**Table 1**Patients Excluded After Screening Assessment (n = 6,849)<sup>a</sup>

<b>Reason for Exclusion</b>	<b>n</b>	<b>%</b>
<b>Signs or symptoms of disc herniation</b>		
Disc Extrusion/ Protrusion/ Bulge	447	6.5
Pain radiating below knee	163	2.4
Lumbar radiculopathy	149	2.2
Leg pain > back pain	78	1.1
Nerve root impingement	6	0.1
Abnormal electrodiagnostic test	5	0.1
<b>Discogenic back pain not confirmed</b>		
Not confirmed with imaging	204	3.0
Not discogenic by MD assessment	115	1.7
Not referred by MD for study for other reasons	16	0.2
<b>Primary problem not related to lumbar spine</b>		
Neck	827	12.1
Problem not back (hip, leg, ankle)	212	3.1
Thoracic spine	207	3.0
Sacral	28	0.4
Shoulder	10	0.1
<b>Diagnosis other than discogenic back pain</b>		
Prior fusion	783	11.4
Fracture	378	5.5
Planned surgery not fusion or disc replacement	421	6.1
Scoliosis	239	3.5
Disc degeneration at more than 2 levels	223	3.3
Lumbar spinal stenosis	147	2.1
Lumbar spondylolisthesis > 25% (grade 1)	86	1.3
Previous multilevel laminectomy	22	0.3
Lesion or cyst	13	0.2
Malignancy or infection	15	0.2
Kyphosis	9	0.1
Developmental deformity	7	0.1
Spondylolysis	6	0.1
Neuropathy	1	<0.1
Instability on flexion-extension x-rays	1	<0.1
<b>Other exclusions</b>		
Age >65 years	922	13.5
Post operative/ wound management	627	9.2
Age <18 years	332	4.8
Did not speak English	77	1.1
Pain duration less than 6 months	25	0.4

<b>Reason for Exclusion</b>	<b>n</b>	<b>%</b>
Medical comorbidity contraindicating surgery	14	0.2
Pregnancy	5	0.1
No phone	1	<0.1

<sup>a</sup>Patients could meet more than one exclusion criterion, but only one is listed here, based on the exclusion criterion first detected during screening.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Baseline comparison of patients who received surgery within 6 months of enrollment and those who did not.

Factor	Level	Non-surgical (n=409)	Surgical (n=86)	p-value between groups
<b><u>Demographics</u></b>				
Sex, %	Male	48	45	0.67
	female	52	55	
Education, %	High school or less	27	25	0.62
	Some College	40	46	
	College degree	32	29	
Race, %	White	84	87	0.45
	Other	16	13	
Age, mean (sd) <sup>a</sup> , years		42.7 (9.3)	42.1 (8.7)	0.60
Work status, %	Working full or part time	50	47	0.06
	On leave, unemployed	24	29	
	Homemaker, student, retired	9	15	
	Disabled	17	8	
<b><u>Clinical characteristics</u></b>				
Duration, %	Less than 12 month	17	14	0.67
	1-5 years	47	52	
	5+ year	35	34	
Previous surgery, %	Yes	21	36	0.004
BMI, %	<24.9	36	29	0.51
	25.9-29.9	36	41	
	30.0+	28	29	
Comorbidity in Charlson index, %	Any	36	41	0.43
Smoker, %	Yes	29	21	0.12
Excessive alcohol/drug screen, %	Positive <sup>b</sup>	12	13	0.83
<b><u>Setting</u></b>				
Enrollment site, %	County medical center	37	26	0.08
	Academic affiliate	22	22	
	Private hospital affiliate	41	52	
<b><u>Baseline physical health measures<sup>c</sup></u></b>				
Roland Score, mean (sd)	0-23	15.9 (5.4)	17.7 (4.2)	0.003
SF-36v.2 Physical function, mean (sd)	Norm-based (0-100)	32.9 (10.7)	28.8 (9.2)	<0.001
Overall Pain Rating, mean (sd)	0-10	6.1 (2.3)	6.5 (1.9)	0.10
Back pain bothersomeness, mean (sd)	1-5	4.0 (0.9)	4.4 (0.7)	0.03
Leg pain bothersomeness, mean (sd)	1-5	3.0 (1.4)	3.5 (1.3)	0.01
<b><u>Baseline mental health measures<sup>c</sup></u></b>				
SF-36v.2 Mental Health, mean (sd)	Norm-based (0-100)	42.2 (12.1)	42.2 (12.3)	0.96
Pain Catastrophizing, mean (sd)	Raw score (0-52)	23.2 (13.4)	25.3 (13.3)	0.20

Factor	Level	Non-surgical (n=409)	Surgical (n=86)	p-value between groups
SCL-90 Somatization, mean (sd)	Raw score (0-4)	1.2 (0.6)	1.1 (0.6)	0.80
SCL-90 Depression, mean (sd)	Raw score (0-4)	1.1 (0.9)	1.1 (0.9)	0.58

<sup>a</sup> sd: standard deviation

<sup>b</sup> Positive alcohol/drug screen was “yes” to either of two questions: excessive use within the past year or desire to cut down.

<sup>c</sup> Higher scores indicate worse pain and function on all outcome measures except the SF-36 Physical Function and Mental Health scales, where higher scores indicate better function.



**Table 3**

Therapeutic interventions reported at baseline and during the treatment designation period.

Intervention	Baseline *		Treatment Period **		p-value
	Non-Surgical Group (n=409)		Surgical Group (n=86)		
	%	%	Non-Surgical Group (n=409)	Surgical Group (n=86)	
<i>Surgery</i>	21	36	0	100	NA
<i>Exercise</i>	80	85	73	78	0.37
<i>Activity restriction</i>	74	80	66	88	<0.001
<i>Non-steroidal anti-inflammatory medications</i>	90	91	65	47	0.27
<i>Opioid pain meds</i>	80	88	64	89	<0.001
<i>Physical therapy/occupational therapy</i>	88	95	52	56	0.61
<i>Bed rest</i>	68	73	52	66	0.018
<i>Massage</i>	55	61	33	28	0.33
<i>Brace or corset</i>	42	46	30	50	0.001
<i>Spinal Injections</i>	63	81	29	26	0.66
<i>Pain program</i>	17	29	17	13	0.34
<i>Chiropractic care</i>	45	52	15	10	0.23
<i>Transcutaneous electrical nerve stimulation (TENS)</i>	25	32	15	19	0.39
<i>Ultrasound</i>	34	44	12	20	0.054
<i>Acupuncture</i>	21	27	7	6	0.74
<i>Cognitive-behavioral therapy</i>	5	8	6	4	0.48
<i>Intradiscal electrothermal coagulation (IDET)</i>	8	11	4	3	0.57

Intervention	* Baseline		** Treatment Period		p-value
	Non-Surgical Group (n=409)	Surgical Group (n=86)	Non-Surgical Group (n=409)	Surgical Group (n=86)	
	%	%	%	%	
<i>Number of different non-surgical interventions received during the treatment designation period only</i>					
0			13	8	0.001
1 to 3			41	29	
4 to 6			39	54	
more than 6			7	9	

NA = not applicable

\* Subjects were asked if they received the treatment at any time in the past prior to enrollment.

\*\* The initial six months after enrollment was the time interval during which treatment group designation was determined. Those patients who received surgery during that period comprised the surgical group. Values indicate any endorsement of that treatment at either the three or the six month post-enrollment assessment.

**Table 4**

Success rates at one year following enrollment for various definitions of success.

<b>Criteria for Success at One Year</b>	<b>Non-Surgical Group</b>	<b>Surgical Group</b>	<b>p-value <sup>a</sup></b>
30% improvement (from baseline) in pain intensity	35%	71%	<0.001
30% improvement (from baseline) in Roland score	25%	57%	<0.001
Working (among those for whom work is relevant) <sup>b</sup>	57%	59%	0.92
No opioid pain medications in the past 3 months	51%	47%	0.51
30% improvement in pain intensity & 30% improvement in Roland	19%	51%	<0.001
30% improvement in pain intensity & 30% improvement in Roland & Working at 12-months <sup>b</sup>	18%	46%	<0.001
30% improvement in pain intensity & 30% improvement in Roland & Working at 12-months <sup>b</sup> & No opioid pain medications in the past 3 months	15%	33%	<0.001

<sup>a</sup> p-values based on logistic regression controlling for age, sex, education, previous surgery, alcohol use, smoking, bmi, race, work status, Charlson comorbidity index, overall pain intensity, back pain bothersome, leg pain bothersome, symptom duration, study recruitment site.

<sup>b</sup> Excluding those patients who at baseline reported being a student, homemaker, retired, or on permanent disability.

**Table 5**

Linear mixed models to estimate treatment effects on primary and secondary outcomes for subjects who received surgery compared to those who received only non-surgical treatment.\*

Outcome	Time point	Non-surgical mean (95% CI)	Surgical mean (95% CI)	P-value
<b>Physical Health Measures</b>				
Roland Disability Index	3 month	15.1 (14.6 – 15.5)	15.3 (14.3 – 16.3)	0.52
	6 month	14.5 (14.0 – 15.0)	12.3 (11.2 – 13.4)	<0.001
	9 month	14.4 (13.9 – 15.0)	9.6 (8.4 – 10.8)	<0.001
	12 month	14.1 (13.5 – 14.8)	8.9 (7.5 – 10.3)	<0.001
Overall Pain Rating	3 month	5.0 (4.8 – 5.2)	5.1 (4.6 – 5.5)	0.43
	6 month	5.3 (5.1 – 5.5)	3.8 (3.3 – 4.3)	<0.001
	9 month	5.3 (5.1 – 5.6)	3.5 (2.9 – 4.0)	<0.001
	12 month	5.2 (4.9 – 5.4)	3.4 (2.9 – 4.0)	<0.001
SF-36v.2 Physical function	3 month	34.6 (33.8 – 35.3)	35.0 (33.2 – 36.7)	0.38
	6 month	35.0 (34.1 – 35.8)	37.5 (35.5 – 39.4)	0.005
	9 month	34.9 (34.0 – 35.8)	43.8 (41.7 – 45.9)	<0.001
	12 month	35.3 (34.3 – 36.3)	43.3 (40.9 – 45.6)	<0.001
Leg pain bothersomeness	3 month	2.8 (2.7 – 2.9)	2.6 (2.4 – 2.9)	0.24
	6 month	2.7 (2.6 – 2.9)	1.7 (1.5 – 2.0)	<0.001
	9 month	2.8 (2.6 – 2.9)	1.5 (1.2 – 1.8)	<0.001
	12 month	2.7 (2.5 – 2.8)	2.0 (1.6 – 2.3)	<0.001
Back pain bothersomeness	3 month	3.8 (3.7 – 3.9)	3.5 (3.3 – 3.7)	0.052
	6 month	3.6 (3.5 – 3.7)	2.9 (2.7 – 3.2)	<0.001
	9 month	3.6 (3.5 – 3.8)	2.7 (2.4 – 2.9)	<0.001
	12 month	3.4 (3.3 – 3.6)	2.8 (2.5 – 3.0)	<0.001
<b>Mental Health Measures</b>				
SF-36v.2 Mental Health	3 month	43.9 (42.8 – 45.1)	45.5 (43.1 – 47.9)	0.26
	6 month	43.5 (42.3 – 44.7)	47.0 (44.4 – 49.6)	0.017
	9 month	43.1 (41.8 – 44.3)	50.0 (47.2 – 52.8)	<0.001
	12 month	43.8 (42.4 – 45.1)	49.1 (46.1 – 52.2)	0.002
Pain Catastrophizing	3 month	21.2 (20.1 – 22.4)	18.8 (16.3 – 21.2)	0.087
	6 month	20.7 (18.4 – 23.0)	11.5 (6.2 – 16.8)	0.002
	9 month	20.5 (17.7 – 23.3)	9.6 (2.7 – 16.5)	0.004
	12 month	20.1 (18.9 – 21.4)	12.1 (9.2 – 15.0)	<0.001
SCL-90 Somatization	3 month	1.06 (1.00 – 1.12)	1.15 (1.02 – 1.29)	0.200
	6 month	1.14 (1.01 – 1.27)	0.76 (0.45 – 1.07)	0.025
	9 month	1.27 (1.10 – 1.44)	0.95 (0.50 – 1.40)	0.19
	12 month	1.14 (0.89 – 1.39)	1.05 (0.44 – 1.65)	0.78
SCL-90 Depression	3 month	1.01 (0.92 – 1.09)	0.85 (0.67 – 1.04)	0.13

Outcome	Time point	Non-surgical mean (95% CI)	Surgical mean (95% CI)	P-value
	6 month	1.00 (0.84 – 1.15)	0.81 (0.45 – 1.16)	0.32
	9 month	1.26 (1.07 – 1.45)	0.70 (0.24 – 1.17)	0.029
	12 month	1.10 (1.01 – 1.19)	0.70 (0.50 – 0.90)	<0.001

\* Controlling for age, sex, education, previous surgery, alcohol use, smoking, body mass index, race, work status, Charlson comorbidity index, overall pain intensity, back pain bothersome, leg pain bothersome, symptom duration, study recruitment site.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript