Laminectomy plus Fusion versus Laminectomy Alone for Lumbar Spondylolisthesis


BACKGROUND

The comparative effectiveness of performing instrumented (rigid pedicle screws affixed to titanium alloy rods) lumbar spinal fusion in addition to decompressive laminectomy in patients with symptomatic lumbar grade I degenerative spondylolisthesis with spinal stenosis is unknown.

METHODS

In this randomized, controlled trial, we assigned patients, 50 to 80 years of age, who had stable degenerative spondylolisthesis (degree of spondylolisthesis, 3 to 14 mm) and symptomatic lumbar spinal stenosis to undergo either decompressive laminectomy alone (decompression-alone group) or laminectomy with posterolateral instrumented fusion (fusion group). The primary outcome measure was the change in the physical-component summary score of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36; range, 0 to 100, with higher scores indicating better quality of life) 2 years after surgery. The secondary outcome measure was the score on the Oswestry Disability Index (range, 0 to 100, with higher scores indicating more disability related to back pain). Patients were followed for 4 years.

RESULTS

A total of 66 patients (mean age, 67 years; 80% women) underwent randomization. The rate of follow-up was 89% at 1 year, 86% at 2 years, and 68% at 4 years. The fusion group had a greater increase in SF-36 physical-component summary scores at 2 years after surgery than did the decompression-alone group (15.2 vs. 9.5, for a difference of 5.7; 95% confidence interval, 0.1 to 11.3; P = 0.046). The increases in the SF-36 physical-component summary scores in the fusion group remained greater than those in the decompression-alone group at 3 years and at 4 years (P = 0.02 for both years). With respect to reductions in disability related to back pain, the changes in the Oswestry Disability Index scores at 2 years after surgery did not differ significantly between the study groups (–17.9 in the decompression-alone group and –26.3 in the fusion group, P = 0.06). More blood loss and longer hospital stays occurred in the fusion group than in the decompression-alone group (P<0.001 for both comparisons). The cumulative rate of reoperation was 14% in the fusion group and 34% in the decompression-alone group (P = 0.05).

CONCLUSIONS

Among patients with degenerative grade I spondylolisthesis, the addition of lumbar spinal fusion to laminectomy was associated with slightly greater but clinically meaningful improvement in overall physical health–related quality of life than laminectomy alone. (Funded by the Jean and David Wallace Foundation and others; SLIP ClinicalTrials.gov number, NCT00109213.)
THE INCREASED USE OF THE LUMBAR SPINAL FUSION PROCEDURE IN THE UNITED STATES, along with the wide variation in practice, is attracting interest from multiple stakeholders, including patients, physicians, payers, and policymakers. In a report published in 2014, spinal fusion (465,000 hospital-based procedures in 2011) accounted for the highest aggregate hospital costs ($12.8 billion in 2011) of any surgical procedure performed in U.S. hospitals. The randomized, controlled Spine Patient Outcomes Research Trial (SPORT) showed that surgery was superior to nonoperative care for the management of lumbar degenerative spondylolisthesis. In SPORT, most patients in the surgical group were treated by means of laminectomy with fusion. Herkowitz et al., in a nonrandomized, prospective, comparative study, found that laminectomy with fusion was superior to laminectomy alone; however, to date, there is no class I evidence that laminectomy plus fusion is superior to laminectomy alone for the treatment of degenerative spondylolisthesis. This lack of evidence complicates efforts to guide and standardize practice, such as through dissemination of clinical practice guidelines. A number of prospective studies with at least 5 years of follow-up after surgery have suggested that lumbar decompression without fusion is associated with excellent outcomes.

The hypothesis tested by the Spinal Laminectomy versus Instrumented Pedicle Screw (SLIP) trial was that lumbar laminectomy with instrumented (rigid pedicle screws affixed to titanium alloy rods) fusion would result in greater improvement than that with laminectomy alone in the primary outcome measure — the change in the physical-component summary score of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36; range, 0 to 100, with higher scores indicating better physical health–related quality of life) — at 2 years. Here, we report the results of the primary 2-year outcome of the SLIP trial, as well as the longer-term 3-year and 4-year outcomes.

METHODS

STUDY DESIGN AND OVERSIGHT

In this randomized, controlled trial, patients from five centers were assessed for eligibility during the period from March 2002 through August 2009; the majority (51 patients) were enrolled at one site (for details on the enrollment of the spine, if they had had previous lumbar spinal surgery, or if they had American Society of Anesthesiologists (ASA) class IV or higher disease (with classes ranging from I to VI and higher classes indicating more severe systemic disease).

PATIENTS

All patients with grade I lumbar spondylolisthesis (degree of spondylolisthesis, 3 to 14 mm) with lumbar stenosis and neurogenic claudication with or without lumbar radiculopathy were eligible for inclusion. Patients were excluded if radiography revealed lumbar instability (motion of >3 mm at the level of listhesis, as measured on flexion-extension radiographs of the lumbar spine), if they were judged by the enrolling surgeon to have lumbar instability because of a history of mechanical low back pain with axial loading of the spine, if they had had previous lumbar spinal surgery, or if they had American Society of Anesthesiologists (ASA) class IV or higher disease (with classes ranging from I to VI and higher classes indicating more severe systemic disease).
Patients were screened and enrolled by trial coordinators at each site. A panel of 10 expert spine surgeons was formed to review a brief clinical vignette plus four standardized radiographic and magnetic resonance images for each patient to assess suitability for randomization. This novel approach appeared to increase patient consent to undergo randomization (see the Supplementary Appendix). Radiographic and magnetic resonance images from each patient were reviewed centrally by two neuroradiologists and one neurosurgeon to verify degenerative lumbar canal stenosis with spondylolisthesis without disk herniation. In addition, independent radiologic review of postoperative computed tomographic scans confirmed adherence to the study protocol.

**INTERVENTIONS**

All patients underwent either decompression alone (decompression-alone group) or decompression with posterolateral instrumented fusion (fusion group) at the single level of spondylolisthesis. Decompression was performed by means of a complete laminectomy with partial removal of the medial facet joint. Patients in the fusion group underwent a lumbar laminectomy as well as implantation of pedicle screws and titanium alloy rods across the level of listhesis, with a bone graft harvested from the iliac crest. The SLIP trial did not include the use of bone morphogenetic protein, interbody devices, or minimally invasive techniques for the placement of percutaneous pedicle screws. All the surgeons routinely performed both operations tested in the trial; each of the surgeons had performed at least 100 laminectomies and 100 posterolateral fusions for lumbar spondylolisthesis before joining the SLIP trial.

**OUTCOME MEASURES**

The primary outcome measure was the change in the SF-36 physical-component summary score at 2 years after surgery. The minimal clinically important difference, which was determined on the basis of previous studies, was prespecified to be 5 points. The secondary outcome measure was the change in the disease-specific ODI score (range, 0 to 100, with higher scores indicating more disability related to back pain). The minimal clinically important difference for the ODI was 10 points. Initial clinical assessments were performed during routine outpatient visits at 1.5 months and 3 months by an independent study coordinator who was not aware of the study hypothesis. After 3 months, validated outcome assessment tools (SF-36 and ODI) were mailed to each patient, who then completed and returned them. A study coordinator attempted to contact patients at least three times to improve patient retention. Additional outcome measures that were prespecified in the protocol included operative complications and reoperations. Reoperation was performed at the discretion of the surgeon; patients were contacted annually by independent study coordinators for assessment of the outcomes of reoperation. A prespecified hospital cost analysis was also described in the protocol, although it has not yet been conducted. Although not explicitly described in the protocol, we collected and reported data on estimated blood loss, operative time, and length of stay in the hospital.

**STATISTICAL ANALYSIS**

The sample size was estimated on the basis of a previous prospective pilot study that was performed by the principal investigator in 2004. No data from the pilot study were included in this report. We assumed a standard deviation of 10 for the change in SF-36 physical-component summary score and a 10% rate of loss to follow-up at 2 years. We estimated that with a sample size of 64 patients (32 patients in each randomized group), the study would have 80% power to detect a between-group difference of 7.5 points in the degree of improvement in SF-36 physical-component summary scores, at a two-sided significance level of 0.05.

The strategy for analysis was developed after the trial was completed but before the examination of the data (see the statistical analysis plan, which is available with the protocol). The baseline characteristics of the patients were compared between the groups with the use of independent-sample t-tests for continuous variables, which are presented as means and standard deviations, and the chi-square test or Fisher’s exact test for categorical variables, which are presented as numbers and percentages. Analyses of the primary outcome were performed among all patients who had follow-up assessments, according to their original randomized treatment assignments.

The between-group comparisons of changes...
in SF-36 and ODI scores from baseline were made with the use of mixed-effects models for repeated measures. An unstructured covariance matrix was specified to account for the within-patient correlations of repeatedly measured outcomes. Fixed effects for site, treatment (decompression with fusion vs. decompression alone), time (1.5 months, 3 months, 6 months, 1 year, 2 years, 3 years, and 4 years after randomization), and time-by-treatment interaction were included and were reported as least-squares means and 95% confidence intervals. We computed the robust standard errors and test statistics involving the fixed effects by specifying the EMPIRICAL option within the PROC MIXED procedure (SAS Institute). Comparisons of least-squares means between the treatment groups at each time point and between time points within each treatment group were performed with the use of appropriate contrasts within the mixed-effect models for repeated measures.19

We compared the percentage of patients in the two groups who had a prespecified minimal clinically important difference of 5 points in the SF-36 physical-component summary score at 2 years by fitting a random-intercept logistic-regression model using PROC GLIMMIX (SAS Institute), with adjustment made for the same list of fixed effects. All analyses were performed with SAS software, version 9.4 (SAS Institute).20

**RESULTS**

**PATIENTS**

Figure 1 shows the enrollment, randomization, and follow-up for the SLIP trial. Overall, 130 patients were screened and identified as eligible; 24 eligible patients declined to participate in either treatment group. Among the remaining 106 patients, 66 consented to undergo randomization and 40 declined to undergo randomization because they had a clear preference for a particular surgical strategy, although they agreed to remain in an observation group. One patient who was randomly assigned to the decompression-alone group never had surgery. There were no crossovers from either randomized strategy. A total of 14 patients in the two groups underwent a subsequent reoperation; data from these patients were not censored and were included in the analysis of the primary outcome measure according to the treatment group to which the patient had been randomly assigned.

The mean age of the study population was 67 years, and 80% were women; the age and the preponderance of women are both consistent with findings in previous reports on patients with degenerative spondylolisthesis.2 Baseline characteristics of patients randomly assigned to the decompression-alone group and of those assigned to the fusion group are shown in Table 1. The patients in the fusion group had a mean baseline SF-36 physical-component summary score that was 3.2 points lower than that in the decompression-alone group (P=0.08). Three of the 28 patients in the fusion group for whom data on ASA classification were available (11%), as compared with none of the 33 in the decompression-alone group with available data, had ASA class III disease (P=0.09). The degree of spondylolisthesis was 5.6 mm in the fusion group and 6.5 mm in the decompression-alone group (P=0.10).

**PRIMARY OUTCOME MEASURE**

At 2 years after surgery, patients in the fusion group had a significantly greater increase in the SF-36 physical-component summary score than did those in the decompression-alone group (15.2 points; 95% confidence interval [CI], 10.9 to 19.5; vs. 9.5 points; 95% CI, 5.2 to 13.8) (Table 2). There was a significant between-group difference in the mean treatment effect (i.e., change in SF-36 physical-component summary score from baseline) of 5.7 points (95% CI, 0.1 to 11.3; P=0.046). The magnitude of the difference in treatment effect was sustained longitudinally over the 4 years after surgery (difference at 4 years, 6.7 points; 95% CI, 1.2 to 12.3; P=0.02) (Table 2 and Fig. 2).

Among the patients who were available for the 2-year follow-up, 24 of 28 in the fusion group and 20 of 29 in the decompression-alone group had a prespecified minimal clinically important difference of 5 points in the SF-36 physical-component summary score. According to a random-intercept logistic-regression model, the predicted rate of a minimal clinically important difference of 5 points at the 2-year follow-up was 91.9% (95% CI, 73.1 to 97.9) among patients in the fusion group and 76.1% (95% CI, 49.7 to 91.1) among patients in the decompression-alone group (difference, 15.8 percentage points; 95% CI, −16.0 to 47.6; P=0.18).

No prespecified plan was outlined for the
Figure 1. Eligibility, Randomization, and Follow-up.
adjustment of baseline differences between the groups. For three of the baseline values (SF-36 physical-component summary, ASA class, and degree of spondylolisthesis), there were marginal, nonsignificant differences between the treatment groups, with P values between 0.05
and 0.10 (Table 1). When a separate analysis was performed to adjust for these baseline differences, the fusion group had greater (although not significantly greater) increases in SF-36 physical-component summary scores than did the decompression-alone group at 2 years (3.9 points; 95% CI, −1.5 to 9.4; P = 0.09), 3 years (5.8 points; 95% CI, 1.1 to 13.7; P = 0.02), and 4 years (4.6 points; 95% CI, −0.9 to 10.2; P = 0.10).

SECONDARY OUTCOME MEASURES AND SURGICAL COMPLICATIONS

The differences between the treatment groups in the amelioration of disability related to low back pain, as measured by the change in ODI score at 2, 3, and 4 years, were not significant (Table 2). As shown in Figure 2C, the fusion group had a lower rate of reoperation over the course of 4 years than did the decompression-alone group (14% vs. 34%, P = 0.05). All the reoperations performed in the decompression-alone group were at the index level to address subsequent clinical instability. In contrast, all the reoperations performed in the fusion group were at an adjacent lumbar level (either disk herniation or clinical instability).

Obesity (body-mass index [the weight in kilograms divided by the square of the height in meters] >30) was not a risk factor for reoperation (rates of reoperation were 21% among obese patients and 28% among nonobese patients, P = 0.58). With respect to surgical complications, blood loss, length of stay, and length of procedure were significantly greater in the fusion group than in the decompression-alone group (Table 3).

**Table 2. Changes in SF-36 Physical-Component Summary and ODI Scores from Baseline.*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Change from Baseline</th>
<th>Difference in Change, Fusion vs. Decompression-Alone (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Decompression-Alone Group</td>
<td>Fusion Group</td>
</tr>
<tr>
<td>SF-36 physical-component summary score</td>
<td></td>
<td>34.7</td>
<td>31.5</td>
</tr>
<tr>
<td>Baseline†</td>
<td></td>
<td>1.5 mo</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 mo</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 mo</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 yr</td>
<td>11.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 yr</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 yr</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 yr</td>
<td>7.4</td>
</tr>
<tr>
<td>ODI score</td>
<td></td>
<td>Baseline†</td>
<td>36.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5 mo</td>
<td>−15.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 mo</td>
<td>−17.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 mo</td>
<td>−20.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 yr</td>
<td>−22.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 yr</td>
<td>−17.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 yr</td>
<td>−17.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 yr</td>
<td>−14.7</td>
</tr>
</tbody>
</table>

* Data are presented as least-squares mean values of changes in SF-36 physical-component summary scores and ODI scores from baseline at each follow-up point. Adjustment for multiplicity was not applied. NA denotes not applicable.
† The baseline scores shown are the mean values in the group.

DISCUSSION

The comparative effectiveness of instrumented fusion with laminectomy versus laminectomy...
alone was the clinical issue examined in the SLIP trial, which screened 130 patients for eligibility, enrolled 106 patients, and randomly assigned 66 patients. The primary outcome measure was the change in physical health–related quality of life at 2 years, as measured by the SF-36 physical-component summary score. Although the outcomes did not differ significantly between the treatment groups at 1 year after surgery, the addition of lumbar fusion to laminectomy was associated with significantly greater increases in the SF-36 physical-component summary score at 2, 3, and 4 years after surgery, which suggests a sustained difference between treatments over time. The between-group differences in the increases in SF-36 physical-component summary score were small but clinically meaningful. We did not observe significant between-group differences with respect to reductions in the ODI score, which was the secondary outcome measure of disability related to back pain.

It is generally agreed that mobile degenerative spondylolisthesis with mechanical low back pain causes instability in the lumbar spine and should be treated with decompression plus fusion.22 One question addressed by the SLIP trial was whether a lumbar laminectomy destabilizes the lumbar spine in the context of a nonmobile degenerative spondylolisthesis. In the SLIP trial, degenerative spondylolisthesis after lumbar laminectomy was sufficiently unstable to require reoperation in at least one third of the patients. This rate of reoperation after laminectomy was higher than that reported in other studies but was consistent with the rate of 28% reported in administrative data.
Table 3. Surgical Complications.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Decompression-Alone Group</th>
<th>Fusion Group</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated blood loss</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients with data</td>
<td>34</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Mean — ml</td>
<td>83.4±63.5</td>
<td>513.7±334.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Length of stay in the hospital</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients with data</td>
<td>33</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Mean — days</td>
<td>2.6±0.9</td>
<td>4.2±0.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Duration of operation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients with data</td>
<td>34</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Mean — min</td>
<td>124.4±34.2</td>
<td>289.6±66.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Major complications — no./total (%)</td>
<td>2/35 (6)†</td>
<td>1/31 (3)†</td>
<td>1.0</td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD.
† The complications in the decompression-alone group included wound infection and new neurologic deficit. The complication in the fusion group was pneumonia. All complications were identified within 30 days. Minor complications were not recorded.

from the state of Washington. Larger prospective registry studies might assess the generalizability of the rate of reoperation in this trial. A newer — but relatively untested — less-invasive strategy of unilateral laminotomy with bilateral decompression may offer an advantage over traditional laminectomy because the midline ligamentous structures are preserved, possibly reducing the risk of reoperations for instability.

We have previously reported risk factors for instability after laminectomy for degenerative spondylolisthesis that include disk height, facet angle, and motion on flexion-extension radiographs. Identifying patients whose spines would be likely to remain stable after surgical decompression may reduce the use of the lumbar spinal fusion procedure and may reduce the rate of complications. We observed a reoperation rate of 14% after laminectomy plus fusion, a rate that was similar to that in SPORT (11%) and in a recent study by Brodke et al. (13.3%).

Although SF-36 physical-component summary scores in the fusion group were statistically higher than the scores in the decompression-alone group at 2, 3, and 4 years, the magnitude of the between-group difference was small. Although patients in the decompression-alone group had a significantly higher risk of early reoperation for instability, revision fusion surgery was associated with subsequent better outcome scores than those before the reoperation. Lumbar fusion was significantly associated with more blood loss and longer operative times and therefore might not be appropriate for elderly patients or for patients with certain coexisting conditions, including osteoporosis.

In the SLIP trial, the strategy for performing a fusion included implantation of rigid pedicle screws affixed to titanium alloy rods, with bone graft harvested from the iliac crest. Currently, some surgeons use minimally invasive techniques and use bone-graft extenders or bone morphogenetic protein instead of bone grafts harvested from the iliac crest. In addition, the use of interbody fusion techniques has increased since the SLIP trial was completed. The SPORT spondylolisthesis trial did not identify any one fusion technique as superior to the others; autografts from the iliac crest were used in nearly one third of the cases, and there were no significant differences in the rate of bony fusion or in the rate of reoperation. The most effective method for creating lumbar fusion is not known.

There were marginal, nonsignificant differences in the baseline variables between the patients in the decompression-alone group and those in the fusion group. Some of the differences in the observed outcomes might be attributed to baseline differences rather than to the randomized treatment. Additional studies are important to validate the observations made in the SLIP trial, which might not be generalizable.

The SF-36 physical-component summary has been shown to be a valid, responsive, and reliable tool for the assessment of degenerative lumbar spinal conditions. The assumption we made when calculating the initial sample size estimate, that 10% of the randomly assigned patients would be lost to follow-up, was reasonably accurate at 1 year (11%) and at 2 years (14%); however, by 4 years, 30% of the patients in the initial randomized treatment groups were lost to follow-up. The interpretation of the differences observed at the 3-year and 4-year time points are weakened by the lower rates of follow-up. Future studies will benefit from larger sample sizes that also include valid disease-specific assessments as primary outcomes.

The well-established higher hospital costs of lumbar fusion may suggest that an overall value assessment might favor decompression alone, as
highlighted in a previous study, as well as in another study in this issue of the Journal on the comparative effectiveness of adding fusion to decompression for lumbar spinal stenosis.\textsuperscript{29,30} Future economic analyses would probably include loss of productivity, reoperations, and the use of outpatient health resources to compare these surgical approaches over a longer period.\textsuperscript{31}

In conclusion, we found that lumbar laminectomy plus fusion was associated with a slightly greater but clinically meaningful improvement in physical health–related quality of life than was laminectomy alone at 2, 3, and 4 years after surgery. Supported by research grants from the Jean and David Wallace Foundation (GH 382) and the Greenwich Lumbar Stenosis SLIP Study Fund (GH 384), which was established by Lucinda B. Watson with support from the Stephanie and Lawrence Flinn, Jr. Charitable Trust and James and Elizabeth Li, and by Alan and Jacqueline Stuart, who provided funds to create a Spine Outcomes Research Center at Lahey Hospital and Medical Center to complete the analysis of the SLIP trial results.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Eylln Wasserberger for serving as national study coordinator for the duration of the study period; Jill Curran, M.S., for serving as national data manager; Susan Christopher, R.N., for serving as study manager; Scott J. Sullivan, M.D., and the late Richard H. Velaj, M.D., for serving as independent radiology reviewers; and Peter K. Dempey, M.D., Juan C. Bartolomei, M.D., Lawrence F. Borges, M.D., and Richard P. Schlenk, M.D., for serving as independent members of the spinal experts network.

LAMINECTOMY PLUS FUSION VERSUS LAMINECTOMY ALONE

REFERENCES


15. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study Questionnaire Short Form 36, and pain scales. Spine J 2008;8:968-74.


Copyright © 2016 Massachusetts Medical Society.