Predictors of Oswestry Disability Index Worsening After Lumbar Fusion

JEFFREY L. GUM, MD; LEAH Y. CARREON, MD, MSC; JEFFREY D. STIMAC, MD; STEVEN D. GLASSMAN, MD

abstract

The authors identified patients with an increase in their Oswestry Disability Index (ODI) score after lumbar spine fusion to evaluate whether this is a plausible definition of deterioration and to determine whether any common patient characteristics exist.

A total of 1054 patients who underwent lumbar spinal fusion and had 2-year follow-up data, including the Short Form 36, the ODI, and numeric rating scales for back and leg pain, were identified. Patients with worsening ODI were compared with the remaining cohort. Twenty-eight patients had an absolute increase (worse) in ODI at 1 year postoperatively. Participants with worsening ODI scores included 13 men and 15 women with an average age of 43.3 years; 15 (54%) were smokers. Common medical comorbidities included obesity and hypertension. Complications occurred in 5 (18%) patients and included wound infection, dural tear, and nerve root injury. Pseudarthrosis was common (n=8; 28%). Twenty-one patients required an additional intervention, including epidural injections, fusion revision, and cervical spine surgery.

It is important to have a clear definition of deterioration to better provide informed consent or choice of treatment. Only 28 (2.6%) patients were identified as having an increase in ODI score at 2-year follow-up.
Over the past 2 decades, surgical success has been defined based on more comprehensive standardized patient-reported outcomes. Commonly used, validated, and reliable health-related quality of life (HRQOL) measures include the Medical Outcomes Study, Short Form 36 (SF-36),1,2 Oswestry Disability Index (ODI),3,4 and visual analog or numeric rating scales (NRS) for back and leg pain.5 The definition of improvement has evolved from concepts such as minimal clinically important difference and substantial clinical benefit, which represent patient perceived and noticeable improvement.6,7 In contrast with accepted definitions of improvement, the concept of being worse after lumbar spinal fusion is not well defined. Current descriptions include the occurrence of peri- or postoperative complications, radiographic nonunion, the need for revision surgery, and improvement that was less than what the patient expected.8 In addition, patients commonly inquire about the possibility of being worse after lumbar fusion. The variability in this definition of being worse or clinical deterioration can make this difficult to answer appropriately. It is important for treating physicians to effectively communicate the possibility of being worse after lumbar spinal fusion so patients can make an informed choice of treatment.9

The Oswestry Disability Index is one of the most commonly used HRQOL measures,10 making an absolute increase in ODI scores (i.e., worse score) at an intermediate follow-up interval (i.e, 2 years) a plausible definition of clinical deterioration after lumbar spine fusion. To the current authors’ knowledge, no articles have specifically examined patients who had a worsening in ODI following lumbar spinal fusion to determine the characteristics that would lead to this deterioration. The goals of the current study are to identify a cohort with an absolute increase in ODI at 2-year follow-up, to evaluate whether a worse ODI score is a plausible definition of deterioration, and to examine patient profiles and operative, radiographic, and clinical data to determine whether any common characteristics exist.

**MATERIALS AND METHODS**

A total of 1399 patients who underwent lumbar spinal fusion from a single spine specialty center with prospectively collected, patient-reported outcomes data were evaluated. The data were retrospectively reviewed after institutional review board approval. Patient-reported outcome measures including the ODI, SF-36, and NRS (range, 0 to 10) for back and leg pain are administered as the standard of care preoperatively and at 1 and 2 years postoperatively. Clinically, the ODI is used as a tool to aid surgical decision making. Patients with scores indicating minimal (scores between 0 and 20) or moderate (scores between 20 and 40) disability are treated with less invasive methods. The ODI is also administered during follow-up visits as a measure of treatment effectiveness.

Of the 1399 patients, 1054 (75%) had a minimum 2-year follow-up. No statistically significant differences were found in demographics and baseline outcome measures between the patients with 2-year follow-up data and those lost to follow-up. Patients who had an increase in ODI scores during the preoperative to 2-years postoperative interval were identified (worsening ODI group). Worsening of the ODI scores was defined as a positive net change in score at 2-year follow-up. The worsening ODI group was then compared with patients who had no change or a decrease in ODI during the same 2-year postoperative interval (no change ODI group).

Baseline data evaluated included age, sex, diagnosis, duration of symptoms, comorbidities, body mass index (BMI), surgical history, work status, smoking history, and psychosocial stressors. Diagnoses included were spondylolisthesis, spinal stenosis, degenerative disk disease, and post-diskectomy instability. Decompression was performed for all patients with spondylolisthesis and spinal stenosis. Patients with all other diagnosis had selective decompression performed when necessary. The Charlson Comorbidity Index was used to quantify medical comorbidities.11 Surgical data, including surgical time, estimated blood loss, number of levels fused, and perioperative and postoperative complications, were reviewed.

For each patient in the worsening ODI group, the authors attempted to identify which complaints were driving their clinical deterioration and what etiologic factors were contributing to those complaints. The authors based this on complaints the patient made during office visits as documented in the clinic notes, surgeon opinion, and diagnostic modalities that were used in the evaluation of the postoperative complaints.

The Health Transition Item, an independent item of the SF-36, asks patients to compare their current health with their health 1-year previously. Patient’s responses ranged from much worse, somewhat worse, about the same, somewhat better, and much better. Statistical analyses were performed using SPSS version 17.0 software (SPSS Inc, Chicago Illinois) to compare patients in the worsening ODI group with those in the no change ODI group.

**RESULTS**

Twenty-eight (2.6%) of the 1054 patients who underwent lumbar spine fusion with 2-year follow-up results had an increase (worse score) in ODI (the worsening ODI group). A total of 1026 patients with 2-year follow-up results had either no change or an improvement in their ODI scores (the no change ODI group). The worsening ODI group included 13 (46%) men and 15 (54%) women. The no change ODI group included 390 (38%) men and 636 (62%) women. A relatively higher percentage of men were in the worsening ODI group compared with the no change ODI group.
group; however, this was not statistically significant ($P = .479$). Mean age was 43.3 and 56.6 years in the worsening ODI and no change ODI groups, respectively ($P = .000$). Mean duration of symptoms was 40 months in the worsening ODI group.

A greater proportion of patients in the worsening ODI group (n=15; 54%) were smokers compared with the no change ODI group (n=133; 13%) ($P = .000$) at the time of surgery (Table 1).

Medical comorbidities in the worsening ODI group included obesity (n=8; 28%), hypertension (n=12; 43%), diabetes mellitus (n=3; 11%), chronic lung disease (n=5; 17%), coronary artery disease (n=5; 17%), hypercholesterolemia (n=4; 14%), thyroid disease (n=3; 11%), and ulcer disease (n=2; 7%). The Charlson Comorbidity Index in the worsening ODI group was 2.2 compared with 3.7 in the no change ODI group ($P = .949$). Mean BMI was 26.3 kg/m² in the worsening ODI group and 29.5 kg/m² in the no change ODI group ($P = .872$). Fifteen (54%) patients in the worsening ODI group were receiving worker’s compensation compared with 41 (4%) patients in the no change ODI group (Table 1) ($P = .949$). Nine (32%) patients in the worsening ODI group reported psychosocial stressors, with depression (n=5; 67%) the most commonly reported. Additional psychosocial stressors included divorce (n=1; 11%), addiction (n=1; 11%), anxiety (n=2; 22%), and death in the family (n=1; 11%).

Fifteen (54%) patients in the worsening ODI group had at least 1 prior spine surgery, including discectomy (n=2), posterior lumbar fusion (n=4), decompression (n=4) only, anterior cervical discectomy and fusion (n=4), or insertion of a spinal stimulator (n=4). Lumbar spinal stenosis was the most common diagnosis (n=7; 25%), followed by degenerative disk disease (n=6; 21%), spondylolisthesis (n=6; 21%), post-discectomy instability (n=5, 18%), and nonunion (n=1; 4%). Perioperative complications occurred in 5 (18%) patients in the worsening ODI group: wound infection in 2 (7%), durotomy in 2 (7%), and nerve root injury in 1 (4%). Pseudarthrosis was also common in this group, occurring in 8 (29%) patients. Twenty-one (75%) patients in the worsening ODI group required an additional intervention, including epidural injections (n=11; 39%), fusion revision (n=12; 42%), or cervical spine surgery (n=1; 4%), with some patients requiring 2 or more procedures.

Comparison of HRQOL measures between the worsening ODI and no change ODI groups showed a statistically significant difference between the 2 groups for all outcome measures (Table 2). Mean change in ODI score was 8.4 points (range, 2 to 30 points) and −19.9 (range, 0 to −72 points) for the worsening ODI and no change ODI groups, respectively ($P = .000$). Mean change in SF-36 physical composite summary score was −1.96 and 6.8 in the worsening ODI and no change ODI groups, respectively ($P = .000$), with patients in the worsening ODI group more likely to have a net negative change. Mean change in NRS back pain score was −0.67 (range, −2 to 4) and −3.13 (range, −4 to 10) in the worsening ODI and no change ODI groups, respectively ($P = .000$). Mean change in NRS leg pain score was −0.77 (range, −3 to 5) −2.81 (range, −8 to 10) in the worsening ODI and no change ODI groups, respectively ($P = .001$).

For the worsening ODI group, the authors tried to identify specific symptoms corresponding with the worsening ODI scores. Most often, patients reported back and leg pain (n=18; 64%), followed by back pain only (n=6; 21%), leg pain only (n=2; 7%), and minimal or no back or leg pain (n=2; 7%). All patients who reported back or leg pain also reported neck pain. A diagnostic workup was performed for all 28 patients in the worsening ODI group to determine potential failure etiology. Adjacent disk level degeneration was the most common cause of failure, with pseudarthrosis or nonunion the second most common.

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>ODI Worsening</th>
<th>ODI Improvement</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>28</td>
<td>1026</td>
<td></td>
</tr>
<tr>
<td>Men, No. (%)</td>
<td>13 (46)</td>
<td>390 (38)</td>
<td>.479</td>
</tr>
<tr>
<td>Mean age, y</td>
<td>43.3</td>
<td>56.6</td>
<td>.000</td>
</tr>
<tr>
<td>Smokers, No. (%)</td>
<td>15 (54)</td>
<td>133 (13)</td>
<td>.000</td>
</tr>
<tr>
<td>CCI</td>
<td>2.2</td>
<td>3.7</td>
<td>.949</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.3</td>
<td>29.5</td>
<td>.872</td>
</tr>
<tr>
<td>Worker’s compensation, No. (%)</td>
<td>15 (54)</td>
<td>41 (4)</td>
<td>.000</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; CCI, Charlson Comorbidity Index.

### Table 2

<table>
<thead>
<tr>
<th>HRQOL</th>
<th>ODI Worsening</th>
<th>ODI Improvement</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODI</td>
<td>8.4</td>
<td>−19.9</td>
<td>.000</td>
</tr>
<tr>
<td>SF-36 (PCS)</td>
<td>−1.96</td>
<td>6.8</td>
<td>.000</td>
</tr>
<tr>
<td>NRS (BACK)</td>
<td>−0.67</td>
<td>−3.13</td>
<td>.000</td>
</tr>
<tr>
<td>NRS (LEG)</td>
<td>−0.77</td>
<td>−2.81</td>
<td>.001</td>
</tr>
</tbody>
</table>

Abbreviations: HRQOL, health-related quality of life; NRS, Numeric Rating Scale; ODI, Oswestry Disability Index; SF-36 PCS, Short-Form 36 (physical component summary).
common etiology (n = 12; 42%), followed by pseudarthrosis (n = 8; 29%), recurrent stenosis (n = 6; 21%), and cervical spine disease (n = 2; 7%).

Health Transition Item responses from the SF-36 included patients reporting the following: 39 (4%) felt much worse, 172 (16%) felt somewhat worse, 462 (44%) felt about the same, 210 (20%) felt somewhat better, and 171 (16%) felt much better.

Discussion

It is important to have distinct, agreed upon, definitions of improvement and deterioration for commonly performed surgical procedures, such as lumbar spine fusion. Although HRQOL measures have become the mainstay for outcome appraisal in spinal surgery, they can sometimes be difficult to translate into clinically meaningful changes in a patient’s life, such as being worse then they felt preoperatively postoperatively. Also, during the informed consent process, patients often inquire about the possibility of being worse after lumbar spinal fusion.

The ODI is a commonly used, disease-specific HRQOL measure for low back disability. For the purposes of the current study, the authors defined being worse or deterioration as an absolute increase in ODI at 2-year follow-up. Only 28 (2.6%) of 1054 patients met this definition. This may underestimate the number of patients worse after lumbar spinal fusion because, in practice, it seems that more than 3% of patients verbalize being worse after lumbar spinal fusion at 2-year follow-up. One plausible explanation is that perhaps only 3% of patients are truly worse from baseline and the remaining patients are “about the same” and express dissatisfaction because their improvement was less than expected. Another more likely explanation for this underestimation may be that ODI is insensitive to deterioration.

Currently, no literature supports how many patients truly report being worse after lumbar spine fusion. The Health Transition Item within the SF-36 asks patients at 1-year follow-up to rate their current health status compared with 1 year previously. Although this is inconsistent with the 2-year follow-up outcome data, it is an internal anchor at 1-year follow-up to provide a calculation of how many patients perceive being worse. Within the current cohort, 211 (20%) patients reported being much worse or somewhat worse when directly surveyed.

Although the ODI is well validated and accepted for communicating statistical and patient-perceived improvement, it has not been thoroughly examined for the patient-perceived deterioration. It has been shown to be responsive to deterioration in a nonlinear manner such that a change of patient-perceived improvement is not necessarily similar to patient-perceived deterioration. To the authors’ knowledge, only 1 study has attempted to define a patient-perceived meaningful deterioration using the ODI. The authors used a 5-point global assessment as the external criteria with a small total population of 289 participants, and 46 responded that they were worse postoperatively. The authors concluded that the minimal clinically important difference for deterioration was smaller and interpreted this finding as either “it takes more to become improved” or more likely “the absence of improvement is assessed as deterioration.”

Overall, it is important for outcome measures to be statistically and clinically relevant. It appears that a worse ODI score at 2-year follow-up does not correlate with, and may underestimate, a patient clinically being worse. Thus, interpretation of worsening of the ODI score following treatment should be done with caution. Although the goal of the current study was to explore a potential definition using a worsening ODI score for deterioration, it is clear that further studies are needed to clarify clinical deterioration using the ODI. Determining an alternative measure of clinical deterioration is beyond the scope of the current article.

Several studies have shown that certain surgical pathologies have more consistent improvements in HRQOLs. Although these studies are mostly focused on improvements in outcome scores and not declines, the authors report that lumbar stenosis, degenerative disk disease, or post-diskectomy instability have smaller magnitudes of ODI improvement compared with spondylolisthesis and scoliosis. The current findings are similar. Eighteen (64%) patients in the worsening ODI group had a diagnosis of lumbar stenosis, degenerative disk disease, or post-diskectomy instability. Comparison between the worsening ODI and no change ODI groups revealed a patient demographic and surgical profile consistent with current literature, including a higher prevalence of younger men, obesity, smokers, patients with multiple medical comorbidities, workers compensation, psychosocial stressors, and revision lumbar surgery. It is well documented that perioperative complications typically do not affect 1- and 2-year outcomes, but complications were relatively common (18%) in the ODI worsening group.

Interestingly, 29% of patients in the worsening ODI group developed neck symptoms requiring cervical spine fusion within 2 years following lumbar fusion. It is reasonable to assume that neck symptoms in these patients contributed to their worsening ODI scores. Although the ODI questionnaire is not specifically designed to account for cervical spine symptoms, it is likely that patients had a significant overlap between neck and back symptoms. Also, if the lumbar symptoms were relieved or improved, cervical symptoms may become the primary driver for a less than desired outcome. Patients should be advised of this possibility. Other etiologic causes of worsening ODI were adjacent level degeneration (42%), pseudarthrosis (29%), and recurrent stenosis (21%).

Although a statistically significant difference was found in all 4 outcome measures between the worsening ODI and
no change ODI groups, the mean NRS for back and leg pain within the wors-
ening ODI group still revealed a slight improvement. It might be expected that patients who worsened after lumbar fu-
sion would have declines in all 4 outcome measures, similar to the ODI improve-
ment group showing improvement in all 4 outcome measures. This is explained by the unidimensional measures of NRS for back or leg pain. NRSs are only focused on specific symptomatology, whereas the ODI is multidimensional and more com-
prehensive. In fact, only 10% of the current cohort had a consistent decline in all 4 HRQOL measures, whereas 90% had some inconsistency. This inconsistency between HRQOL measures has been de-
scribed by Copay et al., who found that only 40.5% of their population had a consistent change in all 4 outcome mea-
sures for both improvement and deterio-
ration. Although the ODI is considered a comprehensive measure of outcome after lumbar spine surgery, the ODI in combi-
nation with additional HRQOL measures, such as the SF-36 and the NRS for back and leg pain, provides a more complete assessment of overall patient-perceived outcome.

It is important for clinicians and pa-
tients to have a transparent definition of deterioration to better provide informed consent or choice of treatment. To the authors’ knowledge, this is the first study to spe-
cifically examine an absolute increase in ODI at 2-year follow-up as a definition of clinical deterioration after lumbar spinal fusion. Only 28 (2.6%) of 1054 pa-
tients were identified as having an abso-
lute increase in ODI at 2-year follow-up. This may underestimate deterioration af-
ter lumbar spine fusion and may be more appropriate as a definition of extreme or substantial clinical deterioration. The pro-
file of this cohort is consistent with cur-
rent literature emphasizing careful selec-
tion of appropriate operative indications and optimization of patient-controlled factors, such as smoking status and psy-
chosocial stressors, can improve preoper-
ative patient counseling and surgical out-
come.

REFERENCES

1. Ware JE Jr, Sherbourne CD. The MOS 36-
item short-form health survey (SF-36). I Con-
ceptual framework and item selection. Med
3. Fairbank JC, Pynsent PB. The Oswestry Dis-
6. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the mini-
mum clinically important difference: a re-
7. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY. Defining sub-
8. Imagama S, Kawakami N, Tsuji T, et al. Peri-
operative complications and adverse events af-
after lumbar spinal surgery: evaluation of 1012 op-
9. Copay AG, Martin MM, Subach BR, et al. Assessment of spine surgery outcomes: in-
consistency of change amongst outcome mea-
12. Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability In-
13. Grote M, Brox JI, Vollestad NK. Concur-
18. Bono CM, Lee CK. The influence of subdi-
19. Cobo Soriano J, Sendino Revuelta M, Fab-
20. Carreon LY, Glassman SD, Howard J. Fusion and nonsurgical treatment for symptomatic lumbar degenerative disease: a systematic re-
lications. J Neurosurg Spine. 2007; 6(4):291-
297.


