Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 5: Correlation between radiographic outcome and function

SANJAY S. DHALL, M.D.,1 TANVIR F. CHOUDHRI, M.D.,1 JASON C. ECK, D.O., M.S.,1
MICHAEL W. GROFF, M.D.,1 ZOHER GHOGAWALA, M.D.,2 WILLIAM C. WATTERS III, M.D.,1
ANDREW T. DAILEY, M.D.,2 DANIEL K. RESNICK, M.D.,3 ALOK SHARAN, M.D.,3
PRAVEEN V. MUMMANENI, M.D.,1 JEFFREY C. WANG, M.D.,10 AND MICHAEL G. KAISER, M.D.11

1Department of Neurological Surgery, University of California, San Francisco, California; 2Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, New York, New York; 3Center for Sports Medicine and Orthopaedics, Chattanooga, Tennessee; 4Department of Neurosurgery, Brigham and Women’s Hospital, Boston, Massachusetts; 5Alan and Jacqueline Stuart Spine Research Center, Department of Neurosurgery, Lahey Clinic, Burlington, and Tufts University School of Medicine, Boston, Massachusetts; 6Bone and Joint Clinic of Houston, Houston, Texas; 7Department of Neurosurgery, University of Utah, Salt Lake City, Utah; 8Department of Neurosurgery, University of Wisconsin, Madison, Wisconsin; 9Department of Orthopaedic Surgery, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York; 10Department of Orthopaedic Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California; and 11Department of Neurosurgery, Columbia University, New York, New York

In an effort to diminish pain or progressive instability, due to either the pathological process or as a result of surgical decompression, one of the primary goals of a fusion procedure is to achieve a solid arthrodesis. Assuming that pain and disability result from lost mechanical integrity of the spine, the objective of a fusion across an unstable segment is to eliminate pathological motion and improve clinical outcome. However, conclusive evidence of this correlation, between successful fusion and clinical outcome, remains elusive, and thus the necessity of documenting successful arthrodesis through radiographic analysis remains debatable. Although a definitive cause and effect relationship has not been demonstrated, there is moderate evidence that demonstrates a positive association between radiographic presence of fusion and improved clinical outcome. Due to this growing body of literature, it is recommended that strategies intended to enhance the potential for radiographic fusion are considered when performing a lumbar arthrodesis for degenerative spine disease. (http://thejns.org/doi/abs/10.3171/2014.4.SPINE14268)

KEY WORDS • fusion • lumbar spine • treatment outcomes • practice guidelines

Recommendations

Grade B

When performing lumbar arthrodesis for degenerative lumbar disease, strategies to achieve successful radiographic fusion should be considered, as there appears to be a correlation between successful fusion and improved clinical outcomes.

Rationale

Achieving a solid arthrodesis following a spinal fusion procedure is generally believed to be an important goal; however, the relationship between successful fusion and clinical outcome has not been fully established. Therefore, the utility of exhaustive radiographic testing to determine fusion status may be questioned. The purpose of this review is to examine the literature regarding the relationship between fusion status and clinical outcome after lumbar arthrodesis procedures performed in the treatment of lumbar spinal degenerative disease. Additional information regarding the methodologies and criteria used to evaluate the evidence discussed below is...
located in the *Methodology* section of the first article in this issue (Part 1: Introduction and methodology).11

**Search Criteria**

For this update, a computerized search of the database of the National Library of Medicine between July 2003 and December 2011 was conducted using the search terms “lumbar spine fusion assessment,” “lumbar spine pseudoarthrosis,” or “lumbar spine fusion outcome.” (The spelling “pseudoarthrosis” was used in searching, but searching on this spelling also retrieves publications with the spelling “pseudarthrosis.”) The search was restricted to references in the English language involving humans. This yielded a total of 1076 references. The titles and abstracts of each of these references were reviewed. Papers not concerned with the assessment of postoperative fusion status or those not focused on adult degenerative lumbar disease (for example, papers focused on trauma-related fractures, infection, scoliosis, or ischemic spondylolisthesis) were discarded. Additional articles were obtained from the bibliographies of the selected articles. Fourteen new references were identified that provided either direct or supporting evidence relevant to the radiographic assessment of lumbar fusion status. These were considered in conjunction with the 37 references from the previous search from 1966 to July 2003.16 Reports involving Level III or better medical evidence are listed in Table 1. Supportive data are provided by additional references listed in the bibliography.

**Scientific Foundation**

Achievement of a solid fusion across the treated motion segments is an integral goal of any lumbar fusion procedure performed to treat low-back pain due to lumbar degenerative disease. Therefore, patients who achieve a solid fusion would be expected to have better clinical outcomes compared with those in whom osseous union does not occur (pseudoarthrosis). However, a number of authors have described patients with pseudoarthrosis with favorable clinical outcomes and patients with solid osseous unions who have poor clinical outcomes.32 The radiographic assessment of lumbar fusion status is imperfect and is not without potential downside to the patient (e.g., exposure to ionizing radiation) and society (e.g., health care resource utilization). If the clinical results associated with lumbar fusion procedures do not correlate with radiographic findings, one can question the utility of exhaustive radiographic study to demonstrate fusion. Furthermore, the incorporation of surgical techniques and adjuncts designed to increase radiographic fusion rates may be inappropriate unless a correlation between radiographic and clinical outcomes can be confirmed. The purpose of this document is to review the evidence for and against such a relationship.

A study correlating clinical outcomes with the results of the gold standard for assessment of lumbar fusion status (open surgical exploration) has not been performed. However, studies do exist in which investigators compared various radiographic fusion assessment techniques with clinical outcomes. In total, we noted 10 Level II and III (4 Level II and 6 Level III) studies relating to correlation between clinical and radiographic outcome. Of these, 7 (3 Level II and 4 Level III) studies, showed a positive correlation between successful arthrodesis on radiographs and good clinical outcome. The remaining 3 studies did not show a positive correlation between radiographic fusion and good clinical outcome. We noted another 7 Level IV and V studies, and 5 of them did not show correlation between radiographic fusion and good clinical outcome.4,8–10,14,19

The Level II studies included the studies by Christensen et al. (2002),1 Kornblum et al. (2004),13 Kim et al. (2006),12 and Thalgott et al. (2009).17 In 2002, Christensen and colleagues published a prospective randomized 2-year follow-up study of 148 patients randomized to posterolateral lumbar fusion (PLF) plus pedicle screw fixation or anterior lumbar interbody fusion (ALIF), PLF, and pedicle screw fixation.1 Clinical outcome was assessed using the Dallas Pain Questionnaire (DPQ), the Low Back Pain Rating Scale (LBPR), and a work status survey. The authors found that patients in both treatment groups exhibited highly significant improvements in all 4 categories of quality of life (DPQ) as well as in the back pain and leg pain index (LBPR) compared with their preoperative status. They identified a significant positive relationship between fusion status and functional outcome: patients with successful radiographic fusion did significantly better than those without solid fusions on 3 of 4 subsections of the DPQ (there was also a nonsignificant improvement on the social concerns subsection).

Kornblum et al.13 retrospectively reviewed data from a randomized trial comparing instrumented to uninstrumented posterolateral lumbar fusion, and they looked specifically at the uninstrumented patients. They found that good/excellent outcomes in 86% of the patients with successful fusion versus 56% in those with pseudoarthrosis (p = 0.01), and similarly VAS scores (for both back pain and leg pain) were statistically higher in patients with successful fusion. It is unclear whether outcomes in patients with uninstrumented pseudoarthrosis can be generalized to patients with instrumented pseudoarthrosis. Kim et al.12 randomized a heterogeneous patient population to 1- or 2-level PLF, posterior lumbar interbody fusion (PLIF), or PLIF+PLF. They found that 91% of patients with fusion had superior clinical results as compared with 41% of patients with nonunion. Thalgott et al.17 randomized 50 patients undergoing ALIF with posterior instrumentation to receive either frozen or freeze-dried femoral allograft. In contrast to the previous 2 Level II studies, this study showed no statistically significant difference in ODI and VAS scores between patients with fusion and those with nonunion.

Of the 6 Level III studies, 4 showed a positive correlation between radiographic fusion and good clinical outcome: the 1995 study by Christensen et al.2 and the studies by Zdeblick,20 Wetzel et al.,24 and Djurasovic et al.2 The remaining 2 studies—the study by Penta and Fraser15 and the study by Epstein16—failed to show a correlation. Christensen et al.2 studied 120 consecutive patients who underwent ALIF. Clinical outcome was evaluated 5–13 years after surgery by using the DPQ. At 2 years postop-
TABLE 1: Correlation between radiographic outcome and function: summary of evidence*  

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Level of Evidence</th>
<th>Correlation</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christensen et al., 2002</td>
<td>II Prospective 2-yr follow-up of 148 pts randomized to PLF+PS or ALIF+PLF+PS. Clinical outcome was assessed w/ DPQ, LBPR, &amp; work status survey scales. Both groups showed highly significant improvement in all 4 categories of life quality (DPQ), back pain, &amp; leg pain index (LBPR) compared w/ preop status. The circumferential fusion pts showed a higher posterolateral fusion rate (92%) than the PLF group (80%) (p = 0.04). Circumferential lumbar fusion produced a higher fusion rate w/ tendency toward better functional outcome.</td>
<td>Correlation was found btwn fusion status &amp; functional outcome. Downgraded to Level III because of heterogeneous pt population (isthmic spondylolisthesis, DDD, previous surgeries) &amp; use of static radiographs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Komblum et al., 2004</td>
<td>II Retrospective review of data from randomized trial comparing uninstrumented to instrumented PLF, looking at uninstrumented cohort. Homogeneous population: all had 1-level degenerative spondylolisthesis, w/ good radiological outcome measures &amp; dynamic radiographs. Good/excellent outcomes in 86% of pts w/ fusion vs 56% in pts w/ pseudarthrosis (p = 0.01); pts w/ fusion had significantly better VAS scores for back pain (p = 0.02) &amp; leg pain (p = 0.001).</td>
<td>Correlation found btwn fusion status &amp; better pain scores. Study was graded Level II because it was retrospective.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim et al., 2006</td>
<td>II 184 pts w/ back w/ or w/o leg pain randomized to 1- or 2-level PLF, PLIF, or PLF+PLIF. Heterogeneous population included pts w/ spinal stenosis &amp; isthmic &amp; degenerative spondylolisthesis. Clinical outcome was assessed w/ DPQ, LBPR, &amp; work status survey scales. Both groups showed highly significant improvement in all 4 categories of life quality (DPQ), back pain, &amp; leg pain index (LBPR) compared w/ preop status. The circumferential lumbar fusion produced a higher fusion rate w/ tendency toward better functional outcome.</td>
<td>91% of pts w/ fusion had superior clinical results as compared to 41% of pts w/ nonunion. Downgraded to Level II because of heterogeneous pt population.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalgott et al., 2009</td>
<td>II 50 pts randomized to ALIF + post instrumentation w/ frozen vs freeze-dried FRA. 80% follow-up. Heterogeneous population included pts w/ spinal stenosis &amp; isthmic &amp; degenerative spondylolisthesis. Validated outcome measures &amp; dynamic radiographs. Of 17 pts w/ nonunion, 41% had superior clinical results as compared to 91% of pts w/ fusion.</td>
<td>Higher rates of pseudarthrosis in the freeze-dried allograft group did not correlate w/ worsened clinical outcome. Study was downgraded to Level II because of the heterogeneous pt population.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zdeblick, 1993</td>
<td>III 124 lumbar fusion pts were prospectively studied. Fusion status was determined using AP &amp; flexion-extension radiography at 1 yr. Clinical results were rated as excellent, good, fair, or poor.</td>
<td>Groups w/ higher fusion rates did better. Downgraded to Level III due to heterogeneity of pt population &amp; lack of validated outcomes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christensen et al., 1996</td>
<td>III 120 consecutive pts, w/ clinical outcome evaluated 5–13 yrs postop using DPQ. At 2 yrs postop, radiological outcome was determined by independent observers: 52% complete fusion, 24% questionable fusion, &amp; 24% definitive pseudarthrosis. Pts w/ complete or questionable union had significantly better results than those w/ nonunion (p &lt; 0.01).</td>
<td>DPQ scores correlated well w/ radiological outcome. Downgraded to Level III due to heterogeneity of pt population &amp; use of static radiographs to assess fusion.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penta &amp; Fraser, 1997</td>
<td>III 103 ALIF pts (from a consecutive series of 125) had clinical (LBOS, VAS, MSPQ, ZDS) outcome assessment &amp; 87 pts also had radiographic fusion assessment (AP/lat radiographs w/ or w/o MRI) ~10 yrs postop. 78% rated themselves as having “complete relief” or “a good deal of relief,” but only 34% had excellent or good LBOS. Clinical outcome was not associated w/ the presence of radiological fusion &amp; was not influenced by the compensation status. Psychological disturbance at review &amp; recop, however, was significantly correlated w/ LBOS. Conclusions: ALIF outcome was strongly affected by psychological makeup of pt; however, the negative effect of compensation observed at 2 yrs seems to dissipate w/ time &amp; becomes insignificant at 10 yrs.</td>
<td>Long-term (~10-yr) presence of radiological fusion was not associated w/ clinical outcome. Downgraded to Level III due to heterogeneity of pt population &amp; retrospective nature of study.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wetzel et al., 1999</td>
<td>III 74 consecutive cases of lumbar fusion. Nonvalidated outcome scores on pain relief &amp; medication usage were used. Patients were observed postop at 5 intervals for ~2 yrs (range 24–35 mos, mean 27 mos). Fusion status was based on flexion-extension radiographs in all cases, w/ selective use of other techniques. Overall fusion rate was 61%. At final follow-up, 60% had improved back pain &amp; 70% had improved leg pain. Fusion (r = 3.3, p = 0.010) correlated positively w/ a successful clinical outcome; the presence of pseudarthrosis negatively correlated w/ a successful clinical outcome.</td>
<td>Solid fusion (r = 3.3, p = 0.010) correlated positively w/ successful clinical outcome. Pseudarthrosis was negatively correlated w/ successful clinical outcome. Downgraded to Level III due to heterogeneous pt population &amp; use of nonvalidated outcome scale.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Correlation was downgraded to Level III because of heterogeneous pt population, use of static radiographs, nonvalidated outcome measures, lack of validated outcomes. (continued)
### TABLE 1: Correlation between radiographic outcome and function: summary of evidence (continued)

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Level of Evidence</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epstein, 2008</td>
<td>III</td>
<td>Prospective study of 75 pts undergoing multilevel laminectomy + uninstrumented fusion (mean no. of levels = 2). Heterogeneous: variable no. of levels, half w/ degenerative spondylolisthesis, some had previous surgery, some smokers, age range 44–82 yrs. Used SF-36 &amp; dynamic radiographs; 100% follow-up. Pseudarthrosis in 13 pts, leading to revision surgery in 1. Reports “nearly identical maximum improvement on SF-36” &amp; no correlation btw pseudarthrosis &amp; clinical outcome. The lesser quality prospective study was downgraded to Level III because of the heterogeneous pt population.</td>
<td>Found no correlation btw pseudarthrosis &amp; clinical results. The lesser quality prospective study was downgraded to Level III because of the heterogeneous pt population.</td>
</tr>
<tr>
<td>Djurasovic et al., 2011</td>
<td>III</td>
<td>Cohort of 193 pts (data collected from 3 trials) w/ heterogeneous diagnoses: spondylolisthesis, instability, DDD, ASD, nonunion; all underwent instrumented PLF. Used CT scans instead of radiographs, ODI, medical outcomes study (MOS), SF-36. No loss to follow-up reported. Compared outcomes in pts w/ &amp; w/o fusion; SF-36 &amp; VAS (back &amp; leg) scores did not show significant btw-groups difference at 2 yrs. But 65% of pts w/ fusion achieved MCID on ODI vs 32% of pts w/o fusion (p = 0.004) &amp; similar for SCB on ODI. 67% vs 50% reached SF-36 MCID (p = 0.152) &amp; 63% vs 41% SCB for SF-36 (p = 0.111). Authors concluded that solid fusion contributes to better outcome.</td>
<td>Presence of solid fusion contributes to improved clinical outcome. Study was downgraded to Level III because of the heterogeneous pt population.</td>
</tr>
</tbody>
</table>

* ALIF = anterior lumbar interbody fusion; AP = anteroposterior; ASD = adjacent-segment disease; DDD = degenerative disc disease; DPQ = Dallas Pain Questionnaire; FRA = femoral ring allograft; HNP = herniated nucleus pulposus; LBOS = Low Back Outcome Scale; LBPR = Low Back Pain Rating Scale; MCID = minimum clinically significant difference; MSPQ = Modified Somatic Perception Questionnaire; ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; pt = patient; SCB = substantial clinical benefit threshold; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale; ZDS = Zung Depression Scale.
had excellent or good LBOS scores. The patients’ clinical outcomes could not be correlated with the presence of radiographic fusion. This study also provides Level III medical evidence against a correlation between radiographic fusion status and clinical outcome following lumbar fusion surgery.

Similarly, Epstein1 prospectively studied 75 patients with heterogeneous diagnoses who underwent multilevel decompression and uninstrumented fusion. She reported “nearly identical maximum improvement of SF-36” and that there was not a correlation between radiographic fusion and good clinical outcome.

Summary

There are a total of 10 Level II and III studies regarding this topic. Of these, 7 showed a positive correlation between radiographic presence of fusion and good clinical outcome. Based on the North American Spine Society (NASS) criteria used in the methodology for these guidelines, these are sufficient data to make a Grade B recommendation that strategies that lead to successful radiographic fusion lead to improved clinical outcomes.

Key Issues for Further Investigation

A prospective observational study involving categorization of patients based on multiple validated outcome–derived outcomes and multimodal radiographic outcome assessment would provide Level II medical evidence supporting or refuting the importance of radiographic fusion.

Acknowledgments

We would like to acknowledge the AANS/CNS Joint Guidelines Committee (JGC) for their review, comments, and suggestions, Laura Mitchell, CNS Guidelines Project Manager, for her organizational assistance and Linda O’Dwyer, medical librarian, for assistance with the literature searches. We would also like to acknowledge the following individual JGC members for their contributions throughout the review process: Timothy Ryken, M.D.; Kevin Crockroft, M.D.; Sepideh Amin-Hanjani, M.D.; Steven N. Kalkanis, M.D.; John O’Toole, M.D., M.S.; Steven Casha, M.D., Ph.D.; Aaron Filler, M.D., Ph.D., F.R.C.S.; Daniel Hoh, M.D.; Steven Hwang, M.D.; Todd McCall, M.D.; Jeffrey J. Olson, M.D.; Julie Pilatis, M.D., Ph.D.; Joshua Rosenow, M.D.; and Christopher Winfree, M.D.

Disclosure

Administrative costs of this project were funded by the Congress of Neurological Surgeons and the Joint Section on Disorders of the Spine and Peripheral Nerves of the American Association of Neurological Surgeons and Congress of Neurological Surgeons. No author received payment or honorarium for time devoted to this project. Dr. Ghogawala receives grants from the Patient Centered Outcomes Research Institute (PCORI) and the National Institutes of Health (NIH). Dr. Groff is a consultant for Depuy Spine and EBI Spine. Dr. Mummaneni owns stock in Spinicity and receives honoraria from DePuy Spine and Globus and royalties from DePuy Spine, Quality Medical Publishers, and Thieme Publishing. Dr. Wang owns stock in Bone Biologics, AxioMed, Amedica, CoreSpine, Expanding Orthopedics, Pioneer, Syndicom, VG Innovations, PearlDriver, Flexuspine, Axis, FzioMed, Benvenue, Promethean, Nexgen, ElectroCore, and Surgitech and holds patents with and receives royalties from Biomet, Stryker, SeaSpine, Aesculap, Osprey, Amedica, Synthes, and Alphatec. The authors report no other potential conflicts of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author contributions to the study and manuscript preparation include the following. Acquisition of data: all authors. Analysis and interpretation of data: all authors. Drafting the article: Dhall. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Dhall. Study supervision: Kaiser.

References

17. Thalgott JS, Fogarty ME, Giuffre JM, Christenson SD, Epstein AK, Aprill C: A prospective, randomized, blinded, single-site study to evaluate the clinical and radiographic differences between frozen and freeze-dried allograft when used as part of a circumferential anterior lumbar interbody fusion procedure. Spine (Phila Pa 1976) 34:1251–1256, 2009

Manuscript submitted March 13, 2014. Accepted April 1, 2014. Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14268.
Address correspondence to: Michael G. Kaiser, M.D., Columbia University, Neurological Surgery, The Neurological Institute, 710 W. 168th St., New York, NY 10032. email: mgk7@columbia.edu.