J Neurosurg Spine 21:23–30, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: Radiographic assessment of fusion status

Tanvir F. Choudhri, M.D.,[1] Praveen V. Mummaneni, M.D.,[2] Sanjay S. Dhall, M.D.,[2] Jason C. Eck, D.O., M.S.,[3] Michael W. Groff, M.D.,[4] Zoher Ghogawala, M.D.,[5] William C. Watters III, M.D.,[6] Andrew T. Dailey, M.D.,[7] Daniel K. Resnick, M.D.,[8] Alok Sharan, M.D.,[9] Jeffrey C. Wang, M.D.,[10] and Michael G. Kaiser, M.D.[11]

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

The ability to identify a successful arthrodesis is an essential element in the management of patients undergoing lumbar fusion procedures. The hypothetical gold standard of intraoperative exploration to identify, under direct observation, a solid arthrodesis is an impractical alternative. Therefore, radiographic assessment remains the most viable instrument to evaluate for a successful arthrodesis. Static radiographs, particularly in the presence of instrumentation, are not recommended. In the absence of spinal instrumentation, lack of motion on flexion-extension radiographs is highly suggestive of a successful fusion; however, motion observed at the treated levels does not necessarily predict pseudarthrosis. The degree of motion on dynamic views that would distinguish between a successful arthrodesis and pseudarthrosis has not been clearly defined. Computed tomography with fine-cut axial images and multiplanar views is recommended and appears to be the most sensitive for assessing fusion following instrumented posterolateral and anterior lumbar interbody fusions. For suspected symptomatic pseudarthrosis, a combination of techniques including static and dynamic radiographs as well as CT images is recommended as an option. Lack of facet fusion is considered to be more suggestive of a pseudarthrosis compared with absence of bridging posterolateral bone. Studies exploring additional noninvasive modalities of fusion assessment have demonstrated either poor potential, such as with ^{99m}Tc bone scans, or provide insufficient information to formulate a definitive recommendation.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14267)

Key Words • **lumbar spine** • **fusion** • **diagnostic techniques** • **practice guidelines**

Recommendations

There is no evidence that conflicts with the previous recommendations in the original version of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine."

Grade A

Following lumbar fusion surgery, static lumbar radiographs are not recommended as a stand-alone method to assess fusion status.

Grade B

Following instrumented posterolateral lumbar fusions (PLFs), CT imaging with fine-cut axial and multiplanar reconstruction views is recommended as a method to assess fusion status. When bilateral posterolateral intertransverse bridging bone is observed on CT scans, the presence of solid fusion is strongly suggested. For the determination of pseudarthrosis, the absence of bilateral facet fusion is more suggestive of true nonunion than the absence of PLF.

Following anterior lumbar interbody fusion (ALIF) with cage instrumentation, CT imaging with fine-cut axi-

Abbreviations used in this paper: ALIF = anterior lumbar interbody fusion; AP = anteroposterior; NPV = negative predictive value; PLF = posterolateral lumbar fusion; PLIF = posterior lumbar interbody fusion; PPV = positive predictive value; RSA = roentgen stereophotogrammetric analysis.

al and multiplanar reconstruction views is recommended as a method to assess fusion status. In this setting, the demonstration of bridging bone posterior to the cage (posterior sentinel sign) on CT scans correlates with the presence of solid fusion with a consensus of raters, but intraobserver variability limits the generalizability of a single rater assessment. The presence of bridging bone anterior to the cage (anterior sentinel sign) also correlates with fusion, with higher specificity but lower sensitivity.

Following uninstrumented lumbar fusion surgery, when noninvasive assessment of fusion status is desired, lateral flexion and extension lumbar radiographs are recommended. The lack of significant motion between vertebrae is highly suggestive of successful fusion.

Grade C

Technetium-99 bone scanning is not recommended as a reliable method to assess fusion status following lumbar fusion surgery.

Several radiographic techniques such as static radiography, lateral flexion-extension radiography, and CT imaging, often in combination, are recommended as options for the noninvasive evaluation of suspected symptomatic lumbar pseudarthrosis. However, the sensitivity and specificity of these noninvasive radiographic tests are imperfect. The specific type of fusion and/or instrumentation surgery, patient characteristics, and clinical scenario can influence the choice of modalities.

Rationale

Lumbar fusion procedures are regularly used to help treat pain and other symptoms that can arise from lumbar degenerative disease. Surgeons performing these procedures may use a number of intraoperative and postoperative strategies to try to promote successful fusion. Solid bony fusion can be definitively determined with direct intraoperative assessment during a fusion exploration surgery. However, noninvasive methods of assessing fusion status are clearly more practical. The radiographic fusion rate is an outcome measure frequently cited in studies evaluating lumbar fusion techniques. However, radiographic fusion is not consistently defined throughout the literature. A previous review examined the literature between 1966 and 2003 regarding the ability of various diagnostic techniques to assess fusion status after lumbar fusion surgery for degenerative disease.[22] The purpose of the current review is to reexamine this topic, incorporating the more recent literature.

Search Criteria

For this update, a computerized search of the database of the National Library of Medicine between July 2004 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 1308 references. The titles and abstracts of each of these references were reviewed. Only papers concerned with the assessment of fusion status following arthrodesis procedures for degenerative lumbar disease were included. Additional articles were obtained from the bibliographies of the selected articles, and 17 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 45 references from a previous search of the literature published between 1966 and July 2003, which was conducted using the same search terms.[22] Reports involving Level III or better medical evidence relevant to the primary question are listed in Table 1. Supportive data are provided by additional references listed in the bibliography.

Scientific Foundation

Open surgical exploration is the only method that allows direct inspection of fusion integrity, and therefore it is considered the gold standard of lumbar fusion assessment.[9,11] Surgical exploration, therefore, is an appropriate benchmark to use in establishing the accuracy and predictive value of noninvasive radiographic studies for the assessment of fusion status following lumbar fusion surgery.

Plain Radiographs (Static)

Anteroposterior (AP) and lateral radiographs can demonstrate a continuous bone mass between adjacent vertebral segments following lumbar fusion. Because of the relatively low cost, widespread availability, and long history as a means of assessing fusion, plain spinal radiography remains a common method of assessment of lumbar fusion.[9] However, the limitations of static plain radiography as a reliable test for determining the presence or absence of a solid fusion have been well documented. Brodsky et al. reported a 64% correlation between preoperative plain radiographs and surgical exploration in a retrospective study of 214 lumbar fusion exploration procedures in patients who had undergone prior PLF.[3] Plain radiography had an 89% sensitivity and 60% specificity for predicting solid fusion. Static radiographs interpreted as demonstrating fusion had a positive predictive value (PPV) of 76%. Those predicting pseudarthrosis had a negative predictive value (NPV) of 78%. These data indicate a 0.18 likelihood ratio for a false-positive result (chance of a pseudarthrosis being discovered at exploration when radiography indicates fusion) and a 2.25 likelihood ratio for a false-negative result (chance of a fusion being discovered at exploration when the radiography suggests pseudarthrosis). The study of Brodsky et al. provides Level I evidence regarding the use of plain lumbar radiography compared with open surgical exploration to assess fusion (see Table 1).

In a similar retrospective study of 75 patients, Kant and coworkers found a 68% correlation between static radiography and surgical exploration of lumbar fusion (sensitivity 85%, specificity 62%, PPV 76%, and NPV 54%).[15] The likelihood ratio for a positive result was 0.81, and the

TABLE 1: Radiographic assessment: summary of evidence*

Authors & Year	Level of Evidence	Description	Comment
Blumenthal & Gill, 1993	I	Retrospective study of 49 pts w/ instrumented lumbar fusion who underwent exploration to remove instrumentation. Appears to be consecutive cases though not explicitly stated. AP & lateral radiographs compared w/ op findings, w/ 69% agreement. Accuracy among observers ranged from 57% to 77%. False-positive rate 42%; false-negative rate 29%. Limited accuracy of plain radiographs in assessing fusion status, w/ low validity (large intra- & interobserver variation).	Level I evidence because study used previously developed diagnostic criteria (w/ plain radiographs) in consecutive pts who also received the gold standard (surgical exploration).
Carreon et al., 2007	I	Retrospective study of fine-cut CT w/ multiplanar reconstructions in 93 consecutive pts w/ instrumented PLF who underwent reexploration (163 levels). CT scans evaluated in blinded fashion by 3 spine surgeons for presence or absence of fusion in the rt & lt posterolateral gutters & rt & lt facets. Interobserver variability lower for assessment of PLF status ($\kappa = 0.62$) than for facet fusion status ($\kappa = 0.42$). When PLF was noted bilaterally, solid fusion at surgery was 8.31 times more likely than nonunion. When facet fusion was observed bilaterally, solid fusion at surgery was 2.90 times more likely than nonunion. When unilateral PLF was found, fusion was 5.37 times more likely than nonunion. However, unilateral radiographic assessment of facet fusion was not predictive of fusion at surgery (0.55 likelihood ratio). For predicting nonunion, bilateral radiographic absence of facet fusion was more predictive of pseudarthrosis (5.19 likelihood ratio) than bilateral absence of PLF (2.90 likelihood ratio).	Level I evidence because study used previously developed diagnostic criteria (intertransverse & facet fusion) in consecutive pts who also received the gold standard (surgical exploration). Moderate to good correlation of CT w/ open exploration (κ 0.42–0.62).
Carreon et al., 2008[7]	I	Retrospective study of fine-cut CT w/ multiplanar reconstructions in 49 consecutive pts w/ ALIF who underwent reexploration (69 levels). CT studies were evaluated in blinded fashion by 5 spine surgeons for fusion status & anterior & posterior sentinel signs. For fusion status, sensitivity ranged from 70% to 90% among raters & specificity ranged from 28% to 85% ($\kappa = 0.25$, $p < 0.0001$). Using majority consensus, 67% of cases were classified correctly as fused (93% sensitivity, 46% specificity). Anterior sentinel sign ($\kappa = 0.34$, $p < 0.0001$) showed 20% sensitivity & 92% specificity. Posterior sentinel sign ($\kappa = 0.23$, $p < 0.0001$) showed 67% sensitivity & 79% specificity.	Level I evidence because study used previously developed diagnostic criteria (anterior & posterior sentinel signs) in consecutive pts who also received the gold standard (surgical exploration). Poor reliability of CT to assess fusion status (low κ). Posterior sentinel sign had highest sensitivity & specificity in predicting fusion following ALIF but had poor reliability (lowest κ) & therefore may be less valuable for an individual rater. Assessment of fusion was done via posterior approach for 44 of 49 pts w/ ALIF, raising question whether this less direct mode of exploration is as accurate.
Bohnsack et al., 1999	II	Retrospective study of 42 pts (40 lumbar) on utility of planar ^{99m}Tc bone scintigraphy to assess fusion just before admission for hardware removal. Based on scintigraphy data, pseudarthrosis was suspected in 5 (12%) & confirmed in 4 during surgery (10%)—2 diagnosed & 2 undiagnosed. Accuracy of the method was 88%, sensitivity 50%, specificity 93%, PPV 40%, & NPV 95%. Sensitivity & PPV of bone scintigraphy are low for possible instability after spinal fusion. The method is not sufficient to diagnose pseudarthrosis reliably after arthrodesis.	Based on low sensitivity, bone scan inadequate to diagnose nonunion. Appears to be consecutive cases though not explicitly stated. Level II evidence because using technique w/o previously developed diagnostic criteria.
Brodsky et al., 1991	II (for CT), I (for others)	Retrospective study of 214 explorations to remove internal fixation devices, batteries, or for failed back surgery in 175 pts w/ PLF. Plain radiographs, polytomography, bending films, &/or CT scans were correlated w/ surgical findings. Significant inaccuracy found for all modalities: plain radiographs 36%, polytomograms 41%, bending films 38%, & axial CT 43% noncorrelations. Axial CT had lowest level of inaccuracy (22%), whereas bending films had the highest (38%).	Significant inaccuracy of plain radiographs, polytomograms, bending films, & axial CT in assessing fusion status. Consecutive pts. Limitation for CT because CT imaging has significantly evolved since study. This study used axial images, only w/ much thicker slices than modern CT imaging.

(continued)

TABLE 1: Radiographic assessment: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Comment
Kant et al., 1995	I	Retrospective study of 75 pts w/ instrumented lumbar fusions. Single-blinded examiner reviewed radiographs immediately before hardware removal & fusion exploration: 68% correlation btwn radiographic evaluation & intraop observation. Sensitivity 85%, specificity 62%, PPV 76%, & NPV 54%.	Level I evidence demonstrating limited accuracy of plain radiographs (consecutive cases using previously developed diagnostic criteria).
Laasonen & Soini, 1989	III	Retrospective study of 48 pts w/ persistent pain after lumbar fusion examined using CT (6-mm slices, selective sagittal reconstruction). 157 findings included: fragmentation of fusion mass (16), hairline pseudarthrosis (9), & spinal stenoses (8). The fusion was explored in 20 pts: 21 of 27 main lesions detected by CT were confirmed; 6 CT findings were partially or totally incorrect. 16 (80%) of 20 correlations of CT & fusion assessments. In 2 cases CT suggested nonunion but fusion was solid at surgery, & in 2 cases CT suggested union but pseudarthrosis was found at surgery.	Moderate (80%) accuracy of CT in assessing fusion. Nonconsecutive pts (20 of 48 had reoperations). Level III because nonconsecutive pts, small nos., & used inferior radiographic analysis of fusion.
Larsen et al., 1996	II	Prospective study of 25 consecutive pts w/ lumbar fusion. All had hardware removal & fusion inspection. Studies to rule out pseudarthrosis included plain radiography, flexion-extension radiography, CT, & bone scintigraphy. Each study was evaluated in blinded fashion by radiologist. At exploration, instrumentation was removed & fusion inspected. No statistically significant correlation found btwn radiographic & surgical findings.	Single-observer blinded study demonstrating no significant correlation btwn radiography & exploration. Consecutive cases. Limitation due to small nos.
Jacobson et al., 1997	III	USG results evaluated in 10 pts after posterolateral thoracic or lumbar fusion w/in 1 wk before second-look surgery. 20 sites evaluated for bone graft, solid fusion, clefts, fluid collections, & hardware visibility. USG & surgical findings were compared. In 3 pts, standard radiographs were reviewed before USG; blinded USG evaluation was performed in the remaining 7. USG correctly identified all 10 sites of pseudarthrosis seen intraoperatively. Of 10 sites w/ solid fusion at surgery, USG depicted 6. At 4 sites (in 2 pts), fusion was mistaken for or obscured by hardware. Overall, sensitivity was 100%, specificity 60%, & accuracy 80%.	Level III despite comparison w/ gold standard (surgical exploration) because intraobserver reliability data are lacking & study population does not appear to be consecutive pts undergoing fusion exploration. Rather, it seems to consist of 10 pts who all had been referred for USG testing.
Fogel et al., 2008	III (CT), II (others)	Retrospective study comparing plain radiographs w/ thin-cut helical CT to assess fusion in 90 pts who underwent surgical exploration following PLIF using interbody cage & PLF (mean 27.2 mos from index procedure). All 90 pts (172 fusion levels) evaluated w/ static plain radiographs read by 2 raters in blinded fashion. 54 pts (109 fusion levels) had fine-cut (1-mm) CT studies w/ multiplanar reconstructions evaluated in blinded fashion by 1 of 3 radiologists. Incidence of pseudarthrosis was 2.3% (4 levels). Sensitivity of both CT & plain radiographs was 100%. Specificity was nearly 90% & was not different btwn the 2 radiographic techniques. Authors concluded that following PLIF (w/ cage) & PLF, plain radiographs (static) & helical CT scans performed "very similarly," & that when radiographs show strong evidence for fusion or nonunion, helical CT is unlikely to provide useful new information. However, they acknowledged that the very low no. of pseudarthrosis levels in the study is a limitation.	Limitations of low nos., limited interobserver reliability assessment, & different raters for the 2 studies. Appears to include all 90 pts who had surgical exploration following fusion. Therefore, can be considered consecutive for plain radiographs (Level II evidence because of limitations). 54 (60%) of 90 pts had CT studies, & therefore study was not consecutive for the CT analysis (Level III evidence).

* ALIF = anterior lumbar interbody fusion; NPV = negative predictive value; PLF = posterolateral lumbar fusion; PLIF = posterior lumbar interbody fusion; PPV = positive predictive value; pts = patients; USG = ultrasonography.

likelihood ratio for a negative result was 2.24. This study provides Level I evidence of the limited accuracy of plain radiographs.

Finally, in a study of 49 patients treated with PLF and posterior lumbar interbody fusion (PLIF) with internal fixation, Blumenthal and Gill compared findings on AP and lateral radiographs (interpreted by 2 surgeons and 2 radiologists) with surgical exploration of the fusion mass at the time of reoperation for hardware removal.[1] They reported a 69% agreement between the radiographic diagnosis and surgical findings. The accuracy among the 4 physicians interpreting the radiographs ranged from 57% to 77% (false-positive rate 42%, false-negative rate 29%). These authors concluded that plain radiography has limited accuracy and validity for the assessment of lumbar fusion. Furthermore, they noted significant intra- and interobserver variation, indicating a lack of reliability (κ 0.4–0.7). Their study provides Level I medical evidence indicating that static radiography is only accurate in determining fusion status in roughly two-thirds of cases. Therefore, based on these studies, static AP and lateral radiographs are not recommended as a stand-alone assessment of the presence of a successful arthrodesis after lumbar fusion surgery for degenerative disease.

Flexion-Extension Radiography

In 1948, Cleveland et al. advocated the use of dynamic (flexion-extension) rather than static radiography for the diagnosis of pseudarthrosis following attempted lumbar fusion surgery.[9] Other authors have also suggested that lateral lumbar flexion-extension radiography has utility in lumbar fusion status assessment.[4] There has been disagreement, however, on the number of allowable degrees of motion at the treated (fused) levels for determining the presence or absence of successful bone fusion following surgery.[20]

Brodsky et al. compared the findings of preexploration lumbar flexion-extension radiography with surgical exploration in a retrospective series of 175 patients who underwent reoperation for various indications following instrumented and noninstrumented lumbar fusion.[3] They found a 62% correlation between preoperative flexion-extension radiography and intraoperative findings at exploration (specificity 37%, sensitivity 96%, PPV 70%, and NPV 86%). Their study provides Level I medical evidence that the absence of motion on flexion-extension radiographs is highly suggestive of a solid fusion. The occurrence of some degree of motion at the treated levels, however, does not necessarily indicate a pseudarthrosis.

Computed Tomography

Since its introduction in the 1970s, CT imaging has been used to assess lumbar fusion. Early studies involved axial sequences alone, with resolution far inferior to modern CT technology. Brodsky et al., in a Level II study, reported the use of 6-mm axial slice CT scans; there was a 57% correlation between fusion assessment based on these scans compared with direct surgical exploration in a retrospective series of 175 patients with 214 total operations.[3] In that study, CT imaging demonstrated a sensitivity of 63%, specificity of 86%, PPV of 72%, and an NPV of 81% in the assessment of fusion status. Laasonen and Soini conducted a retrospective review of 20 patients who underwent CT scanning prior to surgical exploration and found an approximate 80% correlation between the CT study–based diagnosis of fusion and intraoperative diagnosis of fusion.[16] Since the publication of these earlier studies, CT imaging technology has advanced. The use of thin-section axial sequences, improved resolution, and multiplanar imaging capability has enhanced the ability of CT scanning to assess lumbar fusion status.

Initial studies with these advanced CT scanning capabilities for lumbar fusion status assessment did not compare the results with the gold standard of direct surgical exploration. Rather, several Level IV or lower studies investigating the utility of CT imaging in lumbar fusion status determination used other radiographic techniques as the comparison group(s). Lang and colleagues found that the addition of thin-slice and multiplanar CT scanning resulted in a higher rate of detection of pseudarthrosis compared with plain radiography.[18] Similarly, Chafetz et al. demonstrated that direct coronal CT scanning may be more sensitive than 2D reconstructed coronal CT images for the detection of pseudarthrosis.[8] Zinreich and colleagues reported that 3D CT reconstruction may be more sensitive than 2D CT reconstruction for the detection of pseudarthrosis.[26] Siambanes and Mather demonstrated that multiplanar CT imaging detected pseudarthrosis in patients who had undergone PLIF, compared with plain radiography, which had suggested a solid fusion.[25] Santos and colleagues examined 32 patients who underwent ALIF with carbon fiber cages.[23] Plain static radiographs were interpreted to demonstrate fusion at 86% of the assessed levels. Flexion-extension lumbar radiography suggested fusion rates ranging from 74% to 96% in this same group of patients, depending on the method used to analyze the radiographs. The addition of thin-section helical CT scanning reduced the radiographic fusion rate to 65%. The authors concluded that CT scanning is more sensitive than static or flexion-extension lumbar radiography for the detection of pseudarthrosis. Shah et al. reached a similar conclusion in their study of 155 patients who underwent PLIF procedures.[24] They found that CT scanning was more sensitive for the detection of abnormalities than plain radiography. These papers are considered to provide Level IV medical evidence on the utility of CT scanning for the diagnosis of pseudarthrosis following attempted lumbar fusion.

More recently, several studies have compared modern CT imaging with open surgical exploration in the assessment of fusion status following instrumented lumbar fusion surgery. Carreon et al. reported a retrospective study of 93 patients with instrumented PLF who had CT imaging (with fine axial cuts and multiplanar reconstructions) prior to open surgical exploration (163 total levels, mean 49 months after initial fusion surgery).[5] The CT studies were evaluated by 3 spine surgeons who were blinded to findings from the fusion exploration. At each level, the raters evaluated for presence or absence of fusion in the right and left posterolateral gutters and right and left facets. The authors found that the interobserver variability was lower for assessment of PLF status (κ = 0.62) than for facet fusion status (κ = 0.42). When PLF was noted bilaterally on CT, the likelihood ratio for solid fusion at surgery was 8.31 times higher than nonunion. When facet fusion was observed bilaterally on CT, the likelihood ratio for solid fusion at surgery was 2.90 times higher than nonunion. When unilateral PLF was found, fusion was 5.37 times more likely than nonunion. However, unilateral radiographic assessment of facet fusion was not predictive of fusion at surgery (0.55 likelihood ratio). For predicting nonunion, bilateral radiographic absence of facet fusion was more predictive of pseudarthrosis (5.19 likelihood ratio) than bilateral absence of PLF (2.90 likelihood ratio). This study provides Level I evidence on the utility of CT imaging to assess fusion status following instrumented PLF.

In a similar study, Carreon et al. reported on 49 patients who had undergone ALIF with cage instrumentation and in whom CT imaging studies (with fine cuts and multiplanar reconstructions) were obtained prior to open surgical exploration (69 levels, mean 22 months after initial fusion surgery).[7] The CT studies were evaluated by 5 spine surgeons who were blinded to findings from fusion exploration. In addition to general assessment of fusion status, anterior and posterior sentinel signs were assessed. For fusion status, sensitivity ranged from 70% to 90% among raters and specificity ranged from 28% to 85% (κ = 0.25, $p < 0.0001$). Using majority consensus, 67% of cases were classified correctly as fused (93% sensitivity, 46% specificity). The anterior sentinel sign (κ = 0.34, $p < 0.0001$) showed 20% sensitivity and 92% specificity. The posterior sentinel sign (κ = 0.23, $p < 0.0001$) showed a 67% sensitivity and 79% specificity. The anterior sentinel sign was more specific (with low numbers) but had low sensitivity. The posterior sentinel sign had better sensitivity but poor reliability related to interobserver variability. This study provides Level I evidence on the utility of CT imaging to assess fusion status following ALIF with cage instrumentation.

Technetium-99m Bone Scan

Technetium-99m bone scanning has also been used to assess fusion status following lumbar arthrodesis surgery. Bohnsack et al. performed a retrospective study of 42 patients who had undergone prior lumbar fusion with internal fixation and who were candidates for reexploration. The authors obtained ^{99m}Tc bone scans before reoperation for hardware removal.[2] The bone scans suggested pseudarthrosis in 5 patients (12%). Pseudarthrosis was found intraoperatively in 4 patients (10%), 2 of which cases were predicted based on the ^{99m}Tc scanning. The accuracy of ^{99m}Tc bone scanning was 88%, with poor sensitivity (50%) but good specificity (93%). The PPV was only 40%, whereas the NPV was 95%. This study provides Level II medical evidence suggesting that ^{99m}Tc bone scanning is not sufficiently reliable to diagnose pseudarthrosis following a lumbar arthrodesis procedure.[2]

Roentgen Stereophotogrammetric Analysis

Roentgen stereophotogrammetric analysis (RSA) is a technique that uses radiopaque 0.8-mm tantalum markers implanted into each vertebral level at the time of surgery (incorporated into the fusion). Johnsson et al. have described the details of the technique.[14] Postoperatively, the patient undergoes computerized radiographic assessment in which two 40° angled roentgen tubes are used. The radiographic imaging is performed with the patient in different positions (for example, supine and upright) to detect movement. The technique assesses the amount of movement between the fused vertebral bodies in multiple planes. The amount of allowable movement that determines fusion versus nonunion, however, is not well defined. This modality has been evaluated in patients at several centers. In a study of 11 patients treated with lumbar fusion, Johnsson and colleagues compared the results of RSA with those of plain radiography at several postoperative time points.[14] In 8 patients in whom plain radiography had demonstrated successful fusion, RSA revealed a progressive decrease in intervertebral movement over time, with achievement of "rigid fusion" within 3–12 months. In a follow-up study, Johnsson et al. performed RSA in 12 patients with lumbar fusion at multiple postoperative time points.[13] Fusion was determined by plain radiography to be present by the end of the study in all patients. For 6 patients, the authors observed gradual reduction in intervertebral movement over time similar to the other study. However, in the other 6 patients negligible movement was observed on assessment 1 month postoperatively. The fact that negligible movement was noted so soon after surgery, when fusion presumably has not yet occurred, is an interesting observation. Pape and associates used RSA in 10 patients following lumbar arthrodesis.[21] Based on RSA criteria, fusion was thought to be present in all patients. This finding was confirmed with open surgical exploration in all cases. Although this report supports the accuracy of the positive correlation between RSA and successful lumbar arthrodesis, because fusion was present in all patients it is not possible to evaluate the utility of RSA in patients with pseudarthrosis.[21]

Other Techniques

Polytomography was used to assess lumbar fusion status in the pre–CT scanning era, but it has been rarely used since the widespread introduction of CT scanning in the 1970s. In their retrospective study of 214 lumbar fusion exploration procedures in patients who had undergone PLF, Brodsky et al. found only a 59% correlation of fusion status between preoperative polytomographs and intraoperative findings (sensitivity 65%, specificity 84%, PPV 79%, and NPV 73%).[3] This single study provides Level I medical evidence that polytomography cannot be reliably used to determine the presence of solid osseous arthrodesis following lumbar fusion procedures for degenerative disease.

The use of MRI studies to assess for pseudarthrosis following lumbar fusion has been explored by several authors. Lang et al. maintained that MRI added unique information in cases involving lumbar fusion procedures.[17] To date, the importance of this information remains unclear. A single report of the use of ultrasonography to evaluate fusion status was also reviewed.[12] Although the results of this study are promising, the ultrasonography technique has not been rigorously evaluated.

Summary

At present, none of the noninvasive radiographic techniques perfectly correlate with open surgical exploration in the detection of solid fusion or pseudarthrosis for patients following lumbar fusion surgery. The assessment of fusion status with static plain radiography is accurate in only approximately two-thirds of patients treated with lumbar fusion when compared with findings from surgical exploration. Therefore, static plain radiography is not recommended as a stand-alone modality of fusion assessment following lumbar fusion procedures. Lateral flexion-extension radiography can be more effective than stand-alone plain radiography in determining fusion status. Lack of motion between fused lumbar segments on lateral flexion-extension views is highly suggestive of a solid fusion in the absence of spinal instrumentation. However, many lumbar spinal fusion procedures are performed with metallic spinal instrumentation, which can interfere with the radiographic assessment of fusion.

Modern CT imaging (with fine-cut axial and multiplanar reconstruction views) appears to be the most effective noninvasive method of determining fusion status following lumbar fusion surgery. CT imaging can detect pseudarthrosis in some patients in whom fusion appeared to be successful based on plain radiographic criteria. Furthermore, CT imaging has proven to be useful in fusion status assessment even in the presence of spinal instrumentation. However, a rigorous prospective comparison of modern CT scanning and surgical exploration has not been performed. Other radiographic techniques have shown some utility as well. The RSA technique is exquisitely sensitive for the detection of motion between vertebral bodies, and the loss of motion between treated vertebral segments does appear to indicate the presence of fusion. This modality, however, is invasive and is not widely available. Furthermore, the sole comparison of RSA with surgical exploration provided only Level III medical evidence supporting the accuracy of RSA. Overall, it is recommended that multiple modalities be considered for the noninvasive evaluation of symptomatic patients with suspected fusion failure, because no radiographic gold standard exists.

Key Issues for Further Investigation

It is understood that routine open surgical exploration to assess fusion status is not practical and that noninvasive methods are clearly preferable. Clinical experience and studies support CT imaging as the leading noninvasive diagnostic study for the evaluation of fusion status following lumbar fusion surgery, because it appears to have superior sensitivity compared with plain radiography for the detection of pseudarthrosis. However, the data supporting CT imaging for this purpose largely come from retrospective and/or nonrandomized studies. A prospective study of CT imaging prior to surgical exploration for instrumentation removal would provide Level I evidence regarding the accuracy of CT studies compared with the gold standard of surgical exploration. If flexion-extension radiographs were obtained in addition prior to exploration, the influence of internal fixation on the accuracy of flexion-extension radiography could also be assessed. Additional developments in image acquisition and processing technology may permit further improvements in the sensitivity and specificity of noninvasive techniques such as CT to detect the presence of solid fusion or pseudarthrosis following lumbar fusion surgery. In particular, further development of ways to minimize imaging artifacts from surgical implants (e.g., rod/screw, plate, or cage instrumentation) would be welcomed.

Additional studies are also needed to clarify which radiographic location of osseous union correlates best with solid fusion, because studies have demonstrated that various sites of radiographic fusion can have different degrees of correlation with overall fusion status.[5,6] Finally, further studies are needed to better understand and reduce the variability of human raters of fusion status on the radiographic studies by establishing and validating objective criteria. Perhaps someday the computerized imaging/processing technology itself could contribute to the determination of fusion status. As noted above, the most effective techniques of noninvasive assessment of fusion status have generally required the use of ionizing radiation (radiographs, CT scans, etc.). It would be ideal to develop noninvasive techniques that do not require ionizing radiation to assess fusion status. Short of that, research can hopefully develop ways to minimize the radiation exposure related to these techniques.

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 Cockroft, M.D.; Sepideh Amin-Hanjani, M.D.; Steven N. Kalkanis, M.D.; John O'Toole, M.D.; 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 Pilitsis, 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, PearlDiver, 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: Choudhri. 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: Choudhri. Study supervision: Kaiser.

References

1. Blumenthal SL, Gill K: Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. **Spine (Phila Pa 1976) 18:**1186–1189, 1993
2. Bohnsack M, Gossé F, Rühmann O, Wenger K: The value of scintigraphy in the diagnosis of pseudarthrosis after spinal fusion surgery. **J Spinal Disord 12:**482–484, 1999
3. Brodsky AE, Kovalsky ES, Khalil MA: Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. **Spine (Phila Pa 1976) 16 (6 Suppl):**S261–S265, 1991
4. Burkus JK, Dorchak JD, Sanders DL: Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. **Spine (Phila Pa 1976) 28:**372–377, 2003
5. Carreon LY, Djurasovic M, Glassman SD, Sailer P: Diagnostic accuracy and reliability of fine-cut CT scans with reconstructions to determine the status of an instrumented posterolateral fusion with surgical exploration as reference standard. **Spine (Phila Pa 1976) 32:**892–895, 2007
6. Carreon LY, Glassman SD, Howard J: Fusion and nonsurgical treatment for symptomatic lumbar degenerative disease: a systematic review of Oswestry Disability Index and MOS Short Form-36 outcomes. **Spine J 8:**747–755, 2008
7. Carreon LY, Glassman SD, Schwender JD, Subach BR, Gornet MF, Ohno S: Reliability and accuracy of fine-cut computed tomography scans to determine the status of anterior interbody fusions with metallic cages. **Spine J 8:**998–1002, 2008
8. Chafetz N, Cann CE, Morris JM, Steinbach LS, Goldberg HI, Ax L: Pseudarthrosis following lumbar fusion: detection by direct coronal CT scanning. **Radiology 162:**803–805, 1987
9. Cleveland M, Bosworth DM, Thompson FR: Pseudarthrosis in the lumbosacral spine. **J Bone Joint Surg Am 30A:**302–312, 1948
10. Fogel GR, Toohey JS, Neidre A, Brantigan JW: Fusion assessment of posterior lumbar interbody fusion using radiolucent cages: X-ray films and helical computed tomography scans compared with surgical exploration of fusion. **Spine J 8:**570–577, 2008
11. Hilibrand AS, Dina TS: The use of diagnostic imaging to assess spinal arthrodesis. **Orthop Clin North Am 29:**591–601, 1998
12. Jacobson JA, Starok M, Pathria MN, Garfin SR: Pseudarthrosis: US evaluation after posterolateral spinal fusion: work in progress. **Radiology 204:**853–858, 1997
13. Johnsson R, Axelsson P, Gunnarsson G, Strömqvist B: Stability of lumbar fusion with transpedicular fixation determined by roentgen stereophotogrammetric analysis. **Spine (Phila Pa 1976) 24:**687–690, 1999
14. Johnsson R, Selvik G, Strömqvist B, Sundén G: Mobility of the lower lumbar spine after posterolateral fusion determined by roentgen stereophotogrammetric analysis. **Spine (Phila Pa 1976) 15:**347–350, 1990
15. Kant AP, Daum WJ, Dean SM, Uchida T: Evaluation of lumbar spine fusion. Plain radiographs versus direct surgical exploration and observation. **Spine (Phila Pa 1976) 20:**2313–2317, 1995
16. Laasonen EM, Soini J: Low-back pain after lumbar fusion. Surgical and computed tomographic analysis. **Spine (Phila Pa 1976) 14:**210–213, 1989
17. Lang P, Chafetz N, Genant HK, Morris JM: Lumbar spinal fusion. Assessment of functional stability with magnetic resonance imaging. **Spine (Phila Pa 1976) 15:**581–588, 1990
18. Lang P, Genant HK, Chafetz N, Steiger P, Morris JM: Three-dimensional computed tomography and multiplanar reformations in the assessment of pseudarthrosis in posterior lumbar fusion patients. **Spine (Phila Pa 1976) 13:**69–75, 1988
19. Larsen JM, Rimoldi RL, Capen DA, Nelson RW, Nagelberg S, Thomas JC Jr: Assessment of pseudarthrosis in pedicle screw fusion: a prospective study comparing plain radiographs, flexion/extension radiographs, CT scanning, and bone scintigraphy with operative findings. **J Spinal Disord 9:**117–120, 1996
20. McAfee PC, Boden SD, Brantigan JW, Fraser RD, Kuslich SD, Oxland TR, et al: Symposium: a critical discrepancy-a criteria of successful arthrodesis following interbody spinal fusions. **Spine (Phila Pa 1976) 26:**320–334, 2001 [Erratum in **Spine (Phila Pa 1976) 26:**1103, 2001]
21. Pape D, Fritsch E, Kelm J, Müller K, Georg T, Kohn D, et al: Lumbosacral stability of consolidated anteroposterior fusion after instrumentation removal determined by roentgen stereophotogrammetric analysis and direct surgical exploration. **Spine (Phila Pa 1976) 27:**269–274, 2002
22. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, et al: Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: radiographic assessment of fusion. **J Neurosurg Spine 2:**653–657, 2005
23. Santos ER, Goss DG, Morcom RK, Fraser RD: Radiologic assessment of interbody fusion using carbon fiber cages. **Spine (Phila Pa 1976) 28:**997–1001, 2003
24. Shah RR, Mohammed S, Saifuddin A, Taylor BA: Comparison of plain radiographs with CT scan to evaluate interbody fusion following the use of titanium interbody cages and transpedicular instrumentation. **Eur Spine J 12:**378–385, 2003
25. Siambanes D, Mather S: Comparison of plain radiographs and CT scans in instrumented posterior lumbar interbody fusion. **Orthopedics 21:**165–167, 1998
26. Zinreich SJ, Long DM, Davis R, Quinn CB, McAfee PC, Wang H: Three-dimensional CT imaging in postsurgical "failed back" syndrome. **J Comput Assist Tomogr 14:**574–580, 1990

Manuscript submitted March 13, 2014.

Accepted April 1, 2014.

Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14267.

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.