J Neurosurg Spine 21:1, 2014

Introduction

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine

Michael W. Groff, M.D.

Department of Neurosurgery, Brigham and Women's Hospital, Boston, Massachusetts

On behalf of the American Association of Neurological Surgeons/Congress of Neurological Surgeons (AANS/CNS) Joint Section on Disorders of the Spine and Peripheral Nerves, it is with distinct pleasure that I introduce the "Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine." The initial version of these guidelines was originally published in the June 2005 issue of the *Journal of Neurosurgery: Spine*.[1] The update presented in this issue of the *Journal of Neurosurgery: Spine* exemplifies the commitment that organized neurosurgery, in cooperation with our orthopedic colleagues, has made to ensure that this vital source of information continues to evolve and reflect the most current evidence on each of the topics covered.

In a very real sense these guidelines are a credit to all clinicians involved in the care of disorders of the spine. That is the body of work on which the literature is based, which in turn is the foundation of these guidelines. This work is also a credit to the established infrastructure created by the Guidelines Committee of the AANS and CNS to facilitate the production and dissemination of evidence-based guidelines. In addition, the successful publication of this material would not have been possible without the assistance provided by the staff at the Journal of Neurosurgery Publishing Group. Their expertise in the peer-review process, editorial guidance, and transmission of information have enhanced the overall quality and impact of this effort. I personally want to thank Dr. James T. Rutka, M.D., Ph.D., Editor-in-Chief of the *Journal of Neurosurgery: Spine*, for his forbearance and attention to this project.

As the literature has evolved, so too has the process of guidelines development. As there is no well-accepted standard protocol for updating guidelines, a significant amount of time and effort was expended to establish the methodology. Consensus among members of the expert panel was achieved in accord with a well-defined methodology to minimize bias, maximize integrity, and create a final product consistent with the highest ideals of evidence-based medicine. Among the most important missions of the AANS/CNS Joint Section on Disorders of the Spine and Peripheral Nerves is the generation of high-quality evidence that can assist both our membership and the spine community at large in providing the highest-quality care for our patients. Just as this work builds on the foundation created by the original 2005 publication, it is anticipated that this document will in time be updated as well. For the moment, however, it reflects an unbiased synthesis of the literature and points toward the quality spine care that we, as clinicians, aspire to provide.

It is an honor to present this Guideline Update on behalf of the AANS/CNS Joint Section on Disorders of the Spine and Peripheral Nerves to the readership of the *Journal of Neurosurgery: Spine*. This update reflects the highest ideals of the section. It is offered to physicians of all levels who seek a greater understanding of the role lumbar fusion can play in the care of patients with degenerative disease of the lumbar spine.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14190)

Disclosure

Dr. Groff is a consultant for DePuy Spine and Biomet Spine.

Reference

1. Heary RF: Introduction to the guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. **J Neurosurg Spine 2:**636, 2005

Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14190.

J Neurosurg Spine 21:2–6, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: Introduction and methodology

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

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

Fusion procedures are an accepted and successful management strategy to alleviate pain and/or neurological symptoms associated with degenerative disease of the lumbar spine. In 2005, the first version of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine" was published in the *Journal of Neurosurgery: Spine*. In an effort to incorporate evidence obtained since the original publication of these guidelines, an expert panel of neurosurgical and orthopedic spine specialists was convened in 2009. Topics reviewed were essentially identical to the original publication. Selected manuscripts from the first iteration of these guidelines as well as relevant publications between 2005 through 2011 were reviewed. Several modifications to the methodology of guideline development were adopted for the current update. In contrast to the 2005 guidelines, a 5-tiered level of evidence strategy was employed, primarily allowing a distinction between lower levels of evidence. The qualitative descriptors (standards/guidelines/options) used in the 2005 recommendations were abandoned and replaced with grades to reflect the strength of medical evidence supporting the recommendation. Recommendations that conflicted with the original publication, if present, were highlighted at the beginning of each chapter. As with the original guideline publication, the intent of this update is to provide a foundation from which an appropriate treatment strategy can be formulated. *(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14257)*

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

Fusion procedures have become a necessary element of the surgeon's armamentarium in the treatment of lumbar degenerative disease. The application of these surgical procedures continues to expand as technological advances facilitate our ability to achieve a solid arthrodesis and our understanding of the pathological and biomechanical aspects of degenerative spine disease improves.

Utilizing national Medicare data from the Dartmouth Atlas Project, Weinstein et al. have identified a steady increase in lumbar fusion surgeries between 1992 and 2003 in patients over the age of 65, from 0.3/1000 to 1.1/1000 enrollees.[6] A 20-fold variation in regional rates among enrollees was also identified, representing the largest regional variation for any surgical procedure. During this interval the annual amount spent for lumbar fusion surgeries rose 500%, to 482 million dollars in 2003.[6] Although Deyo et al. identified a slight decline in the number of lumbar fusion procedures performed among Medicare beneficiaries between 2002 and 2007, the number of complex fusion procedures increased 15-fold, from 1.3 to 19.9 procedures for every 100,000 beneficiaries.[1] With this increase in the complexity of surgery performed, an increase in costs, morbidity, and resource utilization has also been observed. Utilizing the Nationwide Inpatient Sample database, Kalanithi et al. demonstrated a 70% increase in the rate of complications following lumbar fusion in patients over

Abbreviation used in this paper: NASS = North American Spine Society.

65 years of age when compared with patients between 45 and 64 years of age.[4] As a result of this increasing rate of lumbar fusions, expansion of indications, and complexity of surgery, the socioeconomic impact has become more closely scrutinized, requiring that medical evidence justify the application of these procedures.

In 2005, the first iteration of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine" was published in the *Journal of Neurosurgery: Spine*.[5] This comprehensive compendium outlined 16 topics pertaining to the performance of lumbar fusion surgery for degenerative spinal disease, providing 50 recommendations based on a review of the medical literature published between 1966 and 2003. Given the time dependency of a literature review, clinical practice guidelines are evolving documents that require periodic updating as new information and knowledge accumulates. The purpose of the current series, "Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine," is to incorporate the more recent medical evidence that has been published since the original publication and establish new recommendations.

In 2009, an expert panel of neurosurgical and orthopedic spine surgeons was convened, many having participated in the original guidelines effort. All members had experience with clinical guideline development and had completed the evidence-based medicine course developed by the North American Spine Society (NASS). As the current document is to serve as an update, identical topics and search terms were selected from the original guideline publication.

Methodology

The development of evidence-based clinical guidelines is a multistep process, the basis of which has been well described.[3] The current update was constructed through a series of steps, similar to the previous guideline efforts:

1. Selection of topics to study
   a. As this is an update, the same topics from the original guidelines were chosen.
2. Performing a literature search
   a. Searches were limited to English studies investigating human subjects.
3. Collecting relevant studies for review
   a. Searches were reviewed and studies specifically investigating the topic under consideration were chosen.
4. Assessing the quality and strength of the evidence
   a. Modified NASS strategy
5. Formulation of recommendations based on the evidence
   a. Modified NASS strategy
6. Panel review of the evidentiary tables
   a. Consensus method used to establish uniformity of response
7. Submission of guidelines for peer review

As previously stated, the first two steps were based on the topics and search terms used in the original guideline submission.

The literature searches were conducted with the assistance of a librarian who had extensive experience formulating and conducting evidence-based literature searches. Search terms from the original guidelines were used and altered as deemed necessary. Searches of the National Library of Medicine and Cochrane database were conducted from the termination of the original searches, in 2003, through December of 2011. The abstracts were reviewed and all relevant publications were selected for formal assessment. Bibliographies were reviewed from selected publications and appropriate studies selected. The specifics of each search, including the MeSH terms, are described in each chapter.

Topics were assigned to individual panel members, with the primary assignee intended to perform the assessment of evidence and a second panel member intended to review the evidentiary table prior to presentation to the entire panel. Each assignee formulated preliminary recommendations based on their review of the literature. The expert panel completed final determination of the levels of evidence and recommendation grades after reviewing the evidentiary tables.

In an effort to conform to spine guidelines published from other clinical societies, as well as maintain an objective assessment of the evidence, the current panel elected to deviate from the methodology employed in the original guidelines and use the NASS strategy for evidence assessment and recommendation grading (see Tables 1 and 2). As there are no uniformly accepted methods for downgrading evidence, the panel decided to limit downgrading of evidence by no more than one level to avoid excessive subjectivity.

As the current publication is intended to serve as an update of the previous guidelines, the decision was made to include all Level I and II evidence from the original guidelines. A reevaluation of these studies utilizing the NASS strategy was necessary. The panel agreed not to include lower levels of evidence, as these studies were not likely to enhance the updated recommendations.

Quality of Medical Evidence

The foundation for any evidence-based practice guidelines rests on the assessment of medical evidence. The NASS assessment of medical evidence is a 5-tiered strategy that assigns separate levels to "case series" and "expert opinion" (see Table 1). This highlights the major difference between the 3-tiered approach used in the original guideline publication, where the decision was made to combine all lower levels of evidence. This distinction becomes relevant when grading recommendations.

Each study was categorized according to the underlying objective—therapeutic, diagnostic, or prognostic. The initial level of evidence was determined by defining the overall study design. For example, a randomized control trial would start as Level I evidence while a retrospective review could start no higher than Level III. The study's methodology was then analyzed to determine if the nec-

TABLE 1: Levels of medical evidence for primary research topic

Level	Therapeutic Study—Investigating the Effectiveness of Treatment	Diagnostic Study—Investigating the Accuracy of a Diagnostic Test	Prognostic Study—Investigating the Impact That a Baseline Characteristic Has on Disease Outcome	Economic Analysis—Formulating an Economic Model to Determine the Cost Effectiveness of Treatment
I	1. Well-designed RCT w/ appropriate statistical analysis/ reporting a. No major limitations* b. No more than 1 minor limitation* 2. Systematic review of well-designed RCTs w/ consistent findings	1. Evaluation of previously established diagnostic test/ criteria a. Consecutively enrolled patients b. Application of reference "gold" standard 2. Systematic review of Level I studies	1. Well-designed prospective study w/ patient enrollment occurring at same time point in disease process a. At least 80% follow-up at study end point 2. Systematic review of Level I studies	1. Inclusion of sensible/realistic costs & treatment alternatives a. Data derived from multitude of sources b. Multi-way sensitivity analysis performed 2. Systematic review of Level I studies
II	1. Prospective comparative study 2. Systematic review of Level II studies or review of Level I studies w/ inconsistent findings	1. Formulation of diagnostic criteria/test a. Consecutively enrolled patients b. Application of reference "gold" standard 2. Systematic review of Level II studies	1. Retrospective review 2. Study population derived from untreated controls of an RCT 3. Inferior prospective study a. Patient enrolled at different time points b. Less than 80% follow-up 4. Systematic review of Level II studies	1. Inclusion of sensible/realistic costs & treatment alternatives a. Data derived from limited studies b. Multi-way sensitivity analysis performed 2. Systematic review of Level II studies
III	1. Case control studies 2. Retrospective comparative studies 3. Systematic review of Level III studies	1. Study of nonconsecutive patients a. Failure to consistently apply reference "gold" standard	1. Case control study	1. Study analysis based on incomplete costs & failure to consider alternative treatments 2. Systematic review of Level III studies
IV	1. Case series	1. Case control study 2. Utilization of poor reference standard	1. Case series	1. Failure to include sensitivity analysis
V	1. Expert opinion	1. Expert opinion	1. Expert opinion	1. Expert opinion

* See Table 3 for listing of major and minor limitations of study design utilized to determine level of medical evidence. RCT = randomized controlled trial.

essary criteria were fulfilled to maintain the initial level of evidence. These criteria were in part based on the NASS strategy as well as the panel's scientific and clinical experience and are listed in Table 3. Downgrading of therapeutic studies occurred if at least 1 major or 2 minor limitations were identified. For the other study categories, these criteria were considered as well as those specifically outlined in Table 3.

Studies that met all criteria and contained data that would significantly alter current medical practice would be upgraded; however, no study met these criteria. During the panel review of the evidentiary tables, consensus method was used to resolve any disagreement.[2] Ultimately, the panel achieved unanimous agreement for every study evaluated in the evidentiary tables.

Formulation of Treatment Recommendations

The primary investigator for a given topic, prior to the consensus development process, formulated preliminary recommendations. During panel discussions the decision was made as to which studies would serve as the basis for the final recommendations, and these studies were included within the "Scientific Foundation" for a given topic. In general, if high-quality evidence (Level I and II data) was available to formulate a recommendation, lesser-quality evidence was not included. Studies of low quality that conflicted with high-quality evidence were not included in the evidentiary tables, but this discrepancy was mentioned in the "Scientific Foundation."

The expert panel assigned a grade to each recom-

TABLE 2: Recommendation grades

Grade	Definition
A	Good evidence—2 or more Level I studies w/ consistent findings
B	Fair evidence—single Level I study or multiple Level II or III studies w/ consistent findings
C	Poor evidence—single Level II study or multiple Level IV or V studies
I	Insufficient evidence for recommendation—single Level III, IV, or V study; studies of equivalent strength w/ conflicting findings/conclusions

TABLE 3: Classification of study limitations

Major limitations
Failure to provide a power calculation for an RCT
Failure to utilize validated outcomes measures
Heterogeneous patient population
More than 20% patients lost to follow-up
Failure to perform statistical analysis
Crossover rate >20% btwn cohorts
Minor limitations
Inadequate reporting of baseline demographics
Small sample size (<50 patients per treatment group for comparative studies or <50 total enrolled patients for noncomparative studies)
Failure to describe method of randomization
Lack of flow chart following patients' course through study
Failure to account for patients lost to follow-up
Failure to perform independent clinical or radiographic analysis
Utilization of inferior control cohort (e.g., historical control group)
Treatment & control simultaneously applied to same patient
Failure to standardize surgical technique
Inferior radiographic analysis of fusion (e.g., static radiographs)
Clinical &/or radiographic follow-up <1 year

mendation based on the strength of the supporting evidence. Instead of a qualitative description of recommendation grade, as performed in the original guidelines, the expert panel chose to use recommendation grades modified from NASS (see Table 2). The baseline NASS strategy was used, but modifications were included to address instances in which a single study provided evidence for a specific recommendation. The highest-quality recommendation, Grade "A," required 2 or more Level I studies with consistent findings. Fair evidence, either a single Level I study or consistent findings from multiple Level II or III studies, was given a Grade "B" recommendation. Poor-quality evidence would support a Grade "C" recommendation, including either a single Level II study or consistent findings from Level IV or V studies. Recommendations based on a single Level III or lower-level study or studies of equal strength that demonstrated conflicting results were given a Grade "I" designation.

Summary

As greater emphasis is placed on validating the surgical treatments for our patients, particularly with regard to spine surgery, the necessity for evidence-based clinical guidelines is becoming increasing apparent. Given the time dependency of a literature review, all clinical practice guidelines are evolving documents that require periodic updating. As an update, the current publication was intended to build on the foundation established by the original lumbar fusion guidelines. After careful evaluation, the current expert panel felt it necessary to reconsider the methodology of previous guidelines. These changes were incorporated in an effort to perform a more objective evaluation and allow for easier communication among clinicians from other subspecialty organizations.

Although emphasis has recently been placed on evidence-based clinical practice and improving the method of scientific investigation, the panel frequently encountered studies of inferior quality. Despite this limitation, one objective of the current update is to identify areas of future research and stimulate more objective clinical investigation. It is the hope that the well-informed reader will carefully evaluate the "Scientific Foundation" to understand the justification for a given recommendation. As with previous guideline efforts, there is a risk of specialty bias as no nonsurgical stakeholders were involved in the development of this update. Although the potential for this bias exists, considerable effort was taken to try and objectively evaluate the current literature.

These guidelines are not intended to provide rigid treatment algorithms. Instead, it is hoped that this update will serve as a comprehensive review of the current state of the literature and provide the reader with a foundation to formulate an appropriate individualized treatment plan for a given patient. Furthermore it is the intent of any guideline to identify current limitations of the literature and stimulate further investigational research.

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., 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 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: Kaiser. 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: Kaiser. Study supervision: Kaiser.

References

1. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG: Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. **JAMA 303:**1259–1265, 2010
2. Fink A, Kosecoff J, Chassin M, Brook RH: Consensus methods: characteristics and guidelines for use. **Am J Public Health 74:**979–983, 1984
3. Hayward RS, Laupacis A: Initiating, conducting and maintaining guidelines development programs. **CMAJ 148:**507–512, 1993
4. Kalanithi PS, Patil CG, Boakye M: National complication rates and disposition after posterior lumbar fusion for acquired spondylolisthesis. **Spine (Phila Pa 1976) 34:**1963–1969, 2009
5. 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 1: introduction and methodology. **J Neurosurg Spine 2:**637–638, 2005
6. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES: United States' trends and regional variations in lumbar spine surgery: 1992-2003. **Spine (Phila Pa 1976) 31:**2707–2714, 2006

Manuscript submitted March 11, 2014.
Accepted April 1, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14257.

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.

J Neurosurg Spine 21:7–13, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 2: Assessment of functional outcome following lumbar fusion

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

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

Assessment of functional patient-reported outcome following lumbar spinal fusion continues to be essential for comparing the effectiveness of different treatments for patients presenting with degenerative disease of the lumbar spine. When assessing functional outcome in patients being treated with lumbar spinal fusion, a reliable, valid, and responsive outcomes instrument such as the Oswestry Disability Index should be used. The SF-36 and the SF-12 have emerged as dominant measures of general health-related quality of life. Research has established the minimum clinically important difference for major functional outcomes measures, and this should be considered when assessing clinical outcome. The results of recent studies suggest that a patient's pretreatment psychological state is a major independent variable that affects the ability to detect change in functional outcome.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14258)

Key Words • **fusion** • **lumbar spine** • **treatment outcomes** • **practice guidelines**

Recommendations

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

Grade B

It is recommended that when assessing functional outcome in patients treated for low-back pain due to degenerative disease, a reliable, valid, and responsive outcomes instrument, such as the disease-specific Oswestry Disability Index (ODI), be used (Level II evidence).

It is recommended that when assessing general health-related quality of life (HR-QOL) in patients treated for low-back pain due to degenerative disease that a reliable, valid, and responsive outcomes instrument, such as the 36-Item Short Form Health Survey (SF-36), be used (Level II evidence).

Abbreviations used in this paper: BIS = Balanced Inventory for Spinal Disorders; DRI = Disability Rating Index; HR-QOL = health-related quality of life; ICC = intraclass correlation coefficient; LSOQ = Lumbar Spine Outcomes Questionnaire; MCID = minimum clinically important difference; MCS = mental component summary; ODI = Oswestry Disability Index; PCS = physical component summary; RMDQ = Roland-Morris Disability Questionnaire; ROC = receiver operating characteristic; SF-12 = 12-Item Short Form Health Survey; SF-36 = 36-Item Short Form Health Survey; SIP = Sickness Impact Profile; SRS-22 = 22-Item Scoliosis Research Society; VAS = visual analog scale.

It is recommended that the minimum clinically important difference (MCID) be considered when assessing clinical outcome (Level II evidence).

Rationale

The assessment of functional outcome for patients who undergo lumbar fusion surgery continues to be an area of intense clinical interest. The Institute of Medicine has identified low-back pain treatment options as one of the highest priorities for new comparative effectiveness research.[11] In an effort to improve the reporting of outcomes following lumbar fusion, an emphasis has been placed on the implementation of valid, reliable, and objective outcome measures. The majority of these instruments are patient self-assessment questionnaires that report quality of life. They can be divided into 2 groups: those that seek to measure disease-specific outcomes, such as the ODI, and general health surveys, such as the SF-36.

The original Lumbar Fusion Guidelines recommended the utilization of reliable, valid, and responsive instruments to assess clinical outcome following treatment for low-back pain; however, there was insufficient evidence to standardize the utilization of one instrument over another, and multiple options were suggested. Patient satisfaction scales, however, were discouraged unless no alternative was available. Since the publication of the first generation of guidelines, investigators have continued to evaluate the utility of these instruments in the assessment of patients treated for low-back pain. We have assessed functional outcome measures by evaluating the evidence from a diagnostic perspective. That is, measurement of functional outcome would not be expected to improve outcome per se, but rather should allow investigators to "diagnose" any improvement in outcome following treatment.

Search Criteria

A computerized search of the National Library of Medicine database of the literature published between 2004 and 2011 was performed. The following subject headings and configurations yielded 1297 citations: ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ("Treatment Outcome"[MeSH] OR "Patient Satisfaction"[MeSH] OR "functional outcome"[All Fields] OR "functional outcomes"[All Fields] OR "outcome"[title] OR "outcomes"[title]). An additional search using "lumbar spine surgery," "outcomes," and "validation studies" yielded an additional 11 citations. The titles and abstracts of the 1308 articles were reviewed, and 28 clinical series focusing on adult patients who underwent lumbar fusion procedures were selected for analysis. Among the articles reviewed from this search, 10 have been included in the evidentiary table (see Table 1) along with 5 major articles (Level II evidence) from the original Lumbar Fusion Guidelines.[17] These 15 articles form the basis for these recommendations. Two studies focused on the reliability of new outcome measures. Four studies examined the reliability, validity, and responsiveness of a new lumbar spine outcomes measure. Four additional studies focused on the validity of established lumbar spine outcome measures, 1 study examined the responsiveness of a specific outcome measure, and 1 study calculated MCIDs for 4 major lumbar spine outcome measures. Three studies reported major predictors of functional outcome for lumbar spine patients. Among the 15 studies, 14 studies provided Level II and 1 study provided Level III medical evidence regarding functional outcome measures from a diagnostic perspective.

Scientific Foundation

Characteristics of a Functional Outcome Instrument

The criteria that determine whether a functional outcome instrument appropriately measures the response to treatment have not changed since the publication of the original guidelines in 2005.[17] The accuracy of an outcome instrument is dependent on 3 qualities—reliability, validity, and responsiveness.[6,7,13] Reliability is the measure of an instrument's consistency or reproducibility when reporting observations and is described by the following characteristics: interobserver reliability (the degree to which different observers obtain similar results when measuring the same phenomenon), intraobserver reliability (the extent to which the same observer obtains similar results on repeated observations of a fixed characteristic), test-retest reliability (consistency of an instrument between 2 separate time points, similar to intraobserver reliability, except that the characteristic, if clinical, may change with time), and internal consistency (used to describe the extent to which individual test domains correlate with the composite result).[12]

Reliability of an instrument is measured statistically in a variety of ways, depending on the nature of the recording of the observation: the κ statistic measures agreement between observers or observations beyond chance when the measure is in the form of categorical data, phi is used with dichotomous data, and intraclass correlation coefficient (ICC) is used with continuous data (and can be used with categorical data). In addition, the α statistic is used to measure internal consistency—the degree to which individual aspects (called "domains") of an outcome measure correlate with the composite result. A functional outcome measure is considered highly reliable if the κ value is greater than 0.8. A measure is thought to be moderately reliable if the κ value is between 0.6 and 0.8. A κ value of less than 0.6 suggests that the outcome measure is less reliable.[10] The internal reliability (α) is generally measured using the Cronbach α test to determine whether individual domains of a test correlate with the final composite result.[4]

The second criterion used to evaluate a functional outcome measure is validity, the ability to measure the disease-specific properties of interest. More recent literature compares novel functional outcome measures with previously validated instruments to assess validity.[15] Typically, the Pearson product-moment coefficient of correlation (r) is used to examine the congruency between one outcome measure and another, with $r > 0.80$ representing a strong correlation between measures.[14] Newer measures, such as the 22-Item Scoliosis Research Soci-

TABLE 1: Assessment of functional outcome: summary of evidence*

Authors & Year	Level of Evidence	Study Description	Comments
Fairbank et al., 1980	II	25 pts w/ acute LBP & reasonable prognosis were studied at weekly intervals for 3 wks w/ a functional disability survey. The ODI has 10 categories, each w/ 6 responses graded 0–5. A total of 50 points are possible. Test-retest reliability was κ >0.95 ($p < 0.001$) in 22 pts. Over the 3-wk interval, significant improvement was noted clinically & was detected using the ODI. A paired t-test revealed significant improvement on the ODI over 3 wks ($p < 0.05$).	The ODI is a reliable, valid, & responsive measure for detecting changes in LBP & its functional severity.
Roland & Morris, 1983	II	230 pts w/ acute LBP; 193 were studied at 0, 1, & 4 wks after the episode. Test-retest reliability was done in 20 of 230 cases. The construct validity was qualitatively assessed by comparing the functional questionnaire to the pain rating scale. External reliability was κ >0.90 & internal consistency was α >0.80. Construct validity demonstrated that the RMDQ was able to qualitatively detect pts w/ poorer outcomes from acute lumbago; however, no specific analysis was done.	The RMDQ is reliable for assessment of acute LBP.
Deyo, 1986	II	136 pts were examined in a clinic for a chief complaint of LBP. Evaluation was done using SIP & the modified RMDQ (shortened version of SIP) initially & 3 weeks later. Reliability for both scales was κ >0.80 in pts (n = 10) who had no change in pain. For pts who did not resume full activity (n = 47), the reliability was α >0.60. A strong correlation existed btwn the scales (r = 0.85) & btwn the physical dimension of the SIP & the modified RMDQ (r = 0.89). The modified RMDQ correlated less well w/ the psychosocial dimension of the SIP (r = 0.56).	The SIP & modified RMDQ (shorter) are reliable scales for the assessment of LBP that seem to follow the physical dimension of functional disability. The modified RMDQ is less well suited to follow the psychosocial dimension of functional disability.
Salén et al., 1994	II	1445 participants were divided into 3 groups: 1092 volunteer controls, 306 pts w/ axial skeletal pain, & 47 w/ joint pain. Patients were evaluated using the DRI & an FSQ. External reliability for the DRI was κ >0.80. There was a correlation to the FSQ. The DRI was responsive in detecting improvement after joint replacement.	The DRI is a reliable, valid, & responsive measure in pts w/ axial skeletal pain.
Luo et al., 2003	II	Study of 2520 pts w/ LBP; 506 pts assessed over 3–6 months using SF-12. The reliability coefficient for PCS-12 was 0.77 & for MCS-12 it was 0.80. Validity was verified by correlating results w/ back pain intensity & ODI. Responsiveness was verified.	The SF-12 is a reliable, valid, & responsive outcomes instrument for pts w/ LBP.
Bendebba et al., 2007	II	Study of 2539 pts w/ LBP who were assessed pre-treatment & 2 yrs following treatment using a new disease-specific outcomes instrument: the Lumbar Spine Outcomes Questionnaire (LSOQ). The reliability ICCs were >0.8 for LBP, leg pain, functional disability, & physical symptoms other than pain. The validity was verified by comparing the results to ODI & SF-36. Responsiveness was verified.	The LSOQ appears to be reliable, valid, & responsive in pts w/ LBP.
Lee et al., 2008	II	Study of 98 pts w/ cervical & lumbar disorders scheduled for surgery comparing the validity of SF-12 (v 2) to SF-36 (v 2). The PCS & MCS scores correlated strongly btwn SF-12 & SF-36 (r range 0.88–0.97). Except for general health, most of the other subscales correlated strongly (r range 0.81–0.99).	The SF-12 (v 2) is a valid alternative to the SF-36 for pts w/ lumbar spine disorders.
Copay et al., 2008	II	454 of 460 pts who underwent lumbar spinal surgery were evaluated to assess the MCID for ODI, SF-36 PCS, & back & leg pain scales. Outcomes instruments were administered preoperatively & 1 yr postoperatively. MCID values: ODI, 12.8 points; SF-36 PCS, 4.9 points; back pain, 1.2 points; leg pain, 1.6 points.	Using 2 different anchors (HTI of the SF-36 & Satisfaction & Results scales), it was possible to calculate MCID values for ODI, SF-36 PCS, & back & leg pain scales.
Guilfoyle et al., 2009	II	620 pts undergoing spinal surgery studied. Patients were evaluated using SF-36, disease-specific instruments, & VAS scores. There was 88% follow-up at 3 mos & 74% follow-up at 12–60 mos. Strong correlations btwn SF-36 physical function & bodily pain domains & specific disability scales were observed. SF-36 physical function, bodily pain, general health, vitality, & mental health domains were free of floor & ceiling effects.	The SF-36 is valid for assessing functional outcomes following lumbar spinal surgery.

(continued)

TABLE 1: Assessment of functional outcome: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Study Description	Comments
Walsh et al., 2003	II	970 pts w/ spinal disorders w/ complete baseline data & 3-mo follow-up were used to evaluate the responsiveness of the ODI & summary scales of the SF-36. Follow-up rate is not stated. Based on ROC analysis, measures assessing pain were more responsive than those assessing function. The "bodily pain" domain of the SF-36 was the most responsive to worsening symptoms.	The SF-36 is sufficient to measure health status & function for pts w/ back pain. Disease-specific measures (i.e., ODI) might not be necessary when SF-36 is used. Pain scales seem to be the most responsive measures for pts w/ LBP.
Trief et al., 2006	II	115 (72%) of 160 pts from 2 prospective lumbar fusion trials completed preop & 2-yr postop SF-36, ODI, & VAS (pain) assessments. Higher preop MCS scores predicted less back & leg pain after surgery & better postop ODI scores.	Presurgical emotional status (measured using the SF-36 MCS) is a predictor of pain & ODI outcome 2 yrs after lumbar spinal fusion.
Slover et al., 2006	II	3482 pts who recently underwent lumbar spinal surgery completed 3-mo & 1-yr SF-36 & ODI assessments. Follow-up rate is not stated. The average improvement in SF-36 & ODI was smaller in pts w/ psychosocial comorbidities (active compensation, smoking) or medical disorders, including headaches.	SF-36 & ODI are less responsive when significant psychosocial comorbidities or medical comorbidities were present.
Pahl et al., 2006	II	A cross-sectional, observational assessment of 4442 pts w/ spinal disorders was conducted in order to compare the effect of diagnosis on overall health status. Follow-up rate is not stated. Herniated disc w/ radicular pain, lumbar spinal stenosis w/o deformity, degenerative spondylolisthesis, & painful disc degeneration/ spondylosis were all associated w/ negative impact on all 8 subscales of the SF-36.	Younger pts (age <60 yrs) & pts w/ lumbar disc herniation w/ radicular pain had the greatest negative impact on physical health measured using SF-36.
Svensson et al., 2009	II	101 pts were evaluated using the BIS, SF-36, EQ-5D, & ODI, before undergoing surgery. BIS scales showed 80% units more ordered pairs than disordered when compared w/ the other outcomes measures of pain.	The BIS is a valid disease-specific outcome measure for pts w/ back & leg pain. Responsiveness & reliability were not assessed.
Bridwell et al., 2007	III	Multicenter prospective study of 56 pts w/ degenerative lumbar scoliosis to assess the responsiveness of the SRS-22 instrument. Follow-up rate is not stated. The greatest changes observed from preop state to 2 yrs postop were the SRS self-image domain followed by SRS total, SRS pain, & ODI.	The SRS-22 instrument is more responsive than ODI or SF-12 for detecting improved pain, self-image, & function in pts treated surgically for degenerative lumbar scoliosis.

* BIS = Balance Inventory for Spinal Disorders; DRI = Disability Rating Index; FSQ = Functional Status Questionnaire; HR-QOL = health-related quality of life; HTI = Health Transition Item of the SF-36; ICC = intraclass correlation coefficient; LBP = low-back pain; LSOQ = Lumbar Spine Outcomes Questionnaire; MCID = minimum clinically important difference; ODI = Oswestry Disability Index; PCS = physical component summary scale of the SF-36; pt = patient; RMDQ = Roland-Morris Disability Questionnaire; ROC = receiver operating characteristic; SIP = Sickness Impact Profile; SF-12 = 12-Item Short Form Health Survey; SF-36 = 36-Item Short Form Health Survey; SRS-22 = 22-Item Scoliosis Research Society; VAS = visual analog scale.

ety questionnaire (SRS-22), the Balanced Inventory for Spinal Disorders (BIS), and the Lumbar Spine Outcomes Questionnaire (LSOQ) were compared with the ODI and SF-36, since both of these have been shown to be reliable, valid, and responsive for patients with lumbar degenerative diseases who are undergoing lumbar fusion. However, this is not a direct measure of validity.

Finally, a functional outcome instrument must be responsive. The instrument must be able to detect differences in disease severity among populations and should be able to measure the magnitude of treatment effect.

Summary of Literature From Previous Guidelines

Fairbank and colleagues showed that the ODI is a reliable, valid, and responsive measure for detecting changes in low-back pain and its functional severity.[8] Roland and Morris demonstrated the Roland-Morris Disability Questionnaire (RMDQ) is a reliable assessment of acute low-back pain.[18] Deyo showed the Sickness Impact Profile (SIP) and the modified RMDQ are reliable for the assessment of low-back pain, which appears to follow the physical dimension of functional disability.[5] Salén et al. found the Disability Rating Index (DRI) to be a reliable, valid, and responsive measure in patients with axial skeletal pain (see Table 1).[19]

Minimum Clinically Important Difference

The validation of functional outcome measures allows the researcher to confidently select appropriate tools for clinical studies. In order for clinicians to interpret the relevant changes in a particular outcome score, it is important to define the minimum change that is clinically meaningful. Copay and colleagues performed a rigorous study of 460 patients where preoperative and 1-year postoperative scores were obtained in 454 patients with 99% follow-up.[3] The authors determined the MCID for the ODI (12.8 points), SF-36 physical component summary (PCS) (4.9 points), visual analog scale (VAS) for back pain (1.2 points), and VAS for leg pain (1.6 points). The study used robust and validated techniques to provide Level II evidence (see Table 1).[3]

General Health-Related Quality-of-Life Measures

Lee et al. performed a study of 98 patients scheduled for either lumbar or cervical spine surgery and compared the 12-Item Short Form Health Survey (SF-12, version 2) to the SF-36 (version 2).[14] The physical and mental component summary scores strongly correlated between SF-12 and SF-36: r ranged between 0.88 and 0.97. Except for general health, most of the other subscales correlated strongly (r range 0.81–0.99). This study provides Level II evidence that the SF-12 (version 2) is a valid alternative for the SF-36 for patients with lumbar spinal disorders.[14] This is important because of a substantial decrease in the amount of time necessary for eliciting responses on the part of patients by utilization of the SF-12 rather than the SF-36 (see Table 1).

Guilfoyle et al. performed an outcome study of 620 unselected patients who underwent either cervical or lumbar spinal surgery for degenerative disease.[9] The SF-36 was compared with a wide range of disease-specific outcome measures to determine the utility of a general health-related quality of life (HR-QOL) instrument for assessing functional outcome for patients with degenerative spinal diseases. There was excellent early follow-up (88% at 3 months) and a modest loss at long-term follow-up (74% available for follow-up at 1–5 years). The SF-36 physical function, bodily pain, general health, vitality, and mental health domains were free from ceiling or floor effects that would skew the results. In addition, the physical function and bodily pain domains correlated well with validated disease-specific outcome measures. Bodily pain correlated well with VAS arm or leg scores, and the mental health domain correlated well to validated psychological morbidity assessments. The SF-36 physical function and bodily pain domains demonstrated good responsiveness (standard response mean 1.04–1.72 for physical function and bodily pain) following surgery for lumbar disorders. The authors concluded, based on Level II evidence, that the SF-36 was reliable, valid, and responsive for measuring outcome following lumbar spinal surgery (see Table 1).[9]

Walsh et al. assessed outcome at 3 months in 970 patients undergoing a variety of treatments for lumbar degenerative disorders and compared the responsiveness of disease-specific and general health outcome instruments.[23] In this study cohort, 27% of patients underwent surgery, while most were treated with various nonoperative therapies. The authors used a diagnostic test paradigm, the receiver operating characteristic (ROC), for assessing the responsiveness of the different outcome measures. The "gold standard" measure of clinical improvement was physician-patient consensus. Patients did not complete this portion of the assessment 62% of the time, and therefore the level of evidence was downgraded one level for the purposes of establishing recommendations. The bodily pain, physical function, and PCS scores of the SF-36 compared favorably to the ODI. In general, all outcome measures were more responsive for assessing changes in pain than changes in function. The authors provided Level II evidence that the SF-36 is both valid and responsive for assessing lumbar spinal pain and functional outcomes and that it might not be necessary to include disease-specific outcome measures in all studies when using the SF-36 (see Table 1).[23]

Pahl et al. extended the observation that the SF-36 is valid for assessing lumbar spinal disorders by performing a cross-sectional assessment of 4442 patients with spinal problems.[16] The data were generated from the National Spine Network database which consisted of 11,029 patients. The extent of patient follow-up is not stated, and the statistical methods for handling missing data were not discussed. The study's level of evidence was therefore downgraded by one level. These authors found that the impact on patients with lumbar herniated disc with radiculopathy, lumbar stenosis, lumbar degenerative spondylolisthesis, or painful degenerative lumbar spondylotic disc disease was negative in all 8 subscales of the SF-36. Younger patients (< 60 years) and patients with lumbar disc herniation with radiculopathy had the greatest negative impact on physical health as measured by the SF-36.

The authors provided Level II evidence to expand the validity of the SF-36 outcome measure to include patients with lumbar spinal disorders for which surgery is recommended (see Table 1).

Psychosocial Impact on Functional Outcome

Trief et al. explored the effect of a patient's emotional state on functional outcomes following intervention for lumbar spinal disease.[22] In a study comprising 160 patients from 2 separate lumbar fusion prospective trials, the authors obtained follow-up in 115 patients (72%) at 2 years after surgery. They found that the preoperative SF-36 mental component summary (MCS) score was an independent predictor of postoperative ODI score. Specifically, patients with greater emotional morbidity preoperatively had less improvement in ODI following surgery compared with patients with more normal MCS scores (Level II evidence).[22]

Slover et al. made similar observations from a much larger cohort of patients.[20] In a study of 3482 patients who underwent lumbar spinal surgery, the authors found that psychosocial (litigation, chronic headaches, etc.) and medical comorbidities reduced the responsiveness of SF-36 and ODI.[20] The authors' conclusions regarding the effect of psychosocial comorbidities are considered Level II evidence since the rate of follow-up is not stated for this large cohort of patients (see Table 1).

Recently Validated Functional Outcome Measures

It is beyond the scope of the current Guideline Update to provide a comprehensive list of all validated outcomes measures used to evaluate patients with lumbar degenerative diseases. A review of the recent literature, however, did identify 3 relatively novel outcome tools that may prove useful for future outcomes analysis: the Lumbar Spine Outcomes Questionnaire (LSOQ),[1] the Balanced Inventory for Spinal Disorders (BIS),[21] and the Scoliosis Research Society-22 (SRS-22).[2] The LSOQ was found to have an ICC greater than 0.8, was validated by comparing it with the ODI and SF-36 (coefficients of correlation were between 0.7 and 0.9), and was found to be responsive (observed effect sizes ranged from 0.68 to 1.17 for 24-month change scores).[1] These data provide Level II evidence in support of the LSOQ (see Table 1).

The studies evaluating the BIS and SRS-22 were not as comprehensive as those for the LSOQ. The BIS was found to be valid when compared with other outcomes instruments, including the ODI, SF-36, and EQ-5D, but reliability and responsiveness were not reported.[21] The SRS-22 was found to be more responsive than SF-12 or ODI for patients with lumbar degenerative scoliosis who underwent surgical management (Table 1).[2]

Summary

Since the publication of the first generation of lumbar spinal fusion guidelines in 2005, there have been no data that conflict with the previous recommendations. The ODI has emerged as a dominant disease-specific outcome measure. The SF-36 and more recently the SF-12 have emerged as dominant general health outcome measures. In some studies, there are data to suggest that the SF-36 might be sufficient for measuring functional outcome following lumbar spinal fusion because it has demonstrated equivalent responsiveness and validity with disease-specific measures.

More novel outcome measures have been compared with the ODI and the SF-36 to determine their validity and responsiveness. Recent data demonstrate the importance of a patient's pretreatment psychological state as a major independent variable that affects the ability to detect change in functional outcome measures—no surprise to experienced spinal surgeons. Finally, research has established the MCID in major functional outcomes measures, which will enhance the interpretation of these observations. This information will undoubtedly guide future comparative-effectiveness research for lumbar degenerative diseases.

Key Issues for Future Investigation

There is an increasing amount of data suggesting that patient-specific factors, such as pretreatment psychological status, are relevant in the functional outcome assessment following lumbar fusion. Specific diseases are associated with different baseline characteristics that may influence the response depending on the choice of functional outcome measure. The SRS-22, for example, appears to be more responsive than the ODI or the SF-36 for evaluating the results of lumbar spinal fusion in patients with degenerative scoliosis.[2] Establishing whether various functional measures are better suited to assess clinical outcome for a specific degenerative spine disorder will be an important step in the evolution of functional outcome assessment.

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., 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 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: Ghogawala. 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: Ghogawala. Study supervision: Kaiser.

References

1. Bendebba M, Dizerega GS, Long DM: The Lumbar Spine Outcomes Questionnaire: its development and psychometric properties. **Spine J 7:**118–132, 2007
2. Bridwell KH, Berven S, Glassman S, Hamill C, Horton WC III, Lenke LG, et al: Is the SRS-22 instrument responsive to change in adult scoliosis patients having primary spinal deformity surgery? **Spine (Phila Pa 1976) 32:**2220–2225, 2007
3. 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 8:**968–974, 2008
4. Cronbach LJ: Coefficient alpha and the internal structure of tests. **Psychometrika 16:**297–334, 1951
5. Deyo RA: Comparative validity of the sickness impact profile and shorter scales for functional assessment in low-back pain. **Spine (Phila Pa 1976) 11:**951–954, 1986
6. Deyo RA, Andersson G, Bombardier C, Cherkin DC, Keller RB, Lee CK, et al: Outcome measures for studying patients with low back pain. **Spine (Phila Pa 1976) 19 (18 Suppl):** 2032S–2036S, 1994
7. Deyo RA, Diehr P, Patrick DL: Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. **Control Clin Trials 12 (4 Suppl):**142S–158S, 1991
8. Fairbank JC, Couper J, Davies JB, O'Brien JP: The Oswestry low back pain disability questionnaire. **Physiotherapy 66:** 271–273, 1980
9. Guilfoyle MR, Seeley H, Laing RJ: The Short Form 36 health survey in spine disease—validation against condition-specific measures. **Br J Neurosurg 23:**401–405, 2009
10. Hadley MN, Walters BC, Grabb PA, Oyesiku NM, Przybylski GJ, Resnick DK, et al: Guidelines for the management of acute cervical spine and spinal cord injuries. **Clin Neurosurg 49:**407–498, 2002
11. Institute of Medicine: **100 Initial Priority Topics for Comparative Effectiveness Research.** Washington, DC: National Academy Press, 2009 (http://www.iom.edu/~/media/Files/Report%20Files/2009/ComparativeEffectivenessResearchPriorities/Stand%20Alone%20List%20of%20100%20CER%20Priorities%20-%20for%20web.pdf) [Accessed April 17, 2014]
12. Karanicolas PJ, Bhandari M, Kreder H, Moroni A, Richardson M, Walter SD, et al: Evaluating agreement: conducting a reliability study. **J Bone Joint Surg Am 91 (Suppl 3):**99–106, 2009
13. Kopec JA, Esdaile JM: Functional disability scales for back pain. **Spine (Phila Pa 1976) 20:**1943–1949, 1995
14. Lee CE, Browell LM, Jones DL: Measuring health in patients with cervical and lumbosacral spinal disorders: is the 12-item short-form health survey a valid alternative for the 36-item short-form health survey? **Arch Phys Med Rehabil 89:**829–833, 2008
15. Luo X, George ML, Kakouras I, Edwards CL, Pietrobon R, Richardson W, et al: Reliability, validity, and responsiveness of the short form 12-item survey (SF-12) in patients with back pain. **Spine (Phila Pa 1976) 28:**1739–1745, 2003
16. Pahl MA, Brislin B, Boden S, Hilibrand AS, Vaccaro A, Hanscom B, et al: The impact of four common lumbar spine diagnoses upon overall health status. **Spine J 6:**125–130, 2006
17. 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 2: assessment of functional outcome. **J Neurosurg Spine 2:** 639–646, 2005
18. Roland M, Morris R: A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. **Spine (Phila Pa 1976) 8:**141–144, 1983
19. Salén BA, Spangfort EV, Nygren AL, Nordemar R: The Disability Rating Index: an instrument for the assessment of disability in clinical settings. **J Clin Epidemiol 47:**1423–1435, 1994
20. Slover J, Abdu WA, Hanscom B, Weinstein JN: The impact of comorbidities on the change in short-form 36 and oswestry scores following lumbar spine surgery. **Spine (Phila Pa 1976) 31:**1974–1980, 2006
21. Svensson E, Schillberg B, Kling AM, Nyström B: The Balanced Inventory for Spinal Disorders: the validity of a disease specific questionnaire for evaluation of outcomes in patients with various spinal disorders. **Spine (Phila Pa 1976) 34:**1976–1983, 2009
22. Trief PM, Ploutz-Snyder R, Fredrickson BE: Emotional health predicts pain and function after fusion: a prospective multicenter study. **Spine (Phila Pa 1976) 31:**823–830, 2006
23. Walsh TL, Hanscom B, Lurie JD, Weinstein JN: Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. **Spine (Phila Pa 1976) 28:**607–615, 2003

Manuscript submitted March 11, 2014.
Accepted April 1, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14258.

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.

J Neurosurg Spine 21:14–22, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 3: Assessment of economic outcome

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

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

A comprehensive economic analysis generally involves the calculation of indirect and direct health costs from a societal perspective as opposed to simply reporting costs from a hospital or payer perspective. Hospital charges for a surgical procedure must be converted to cost data when performing a cost-effectiveness analysis. Once cost data has been calculated, quality-adjusted life year data from a surgical treatment are calculated by using a preference-based health-related quality-of-life instrument such as the EQ-5D. A recent cost-utility analysis from a single study has demonstrated the long-term (over an 8-year time period) benefits of circumferential fusions over stand-alone posterolateral fusions. In addition, economic analysis from a single study has found that lumbar fusion for selected patients with low-back pain can be recommended from an economic perspective. Recent economic analysis, from a single study, finds that femoral ring allograft might be more cost-effective compared with a specific titanium cage when performing an anterior lumbar interbody fusion plus posterolateral fusion.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14259)

Key Words • **fusion** • **lumbar spine** • **cost-effectiveness** • **outcomes** • **practice guidelines**

Recommendations

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

Grade B

There is Level I evidence (single study) to recommend the use of a circumferential fusion (ALIF + posterolateral fusion) as a more cost-effective option (over an 8-year time period) than stand-alone posterolateral fusion.

There is Level I evidence (single study) to recommend either total lumbar disc replacement (TDR) or lumbar fusion from an economic perspective for the treatment of selected patients with chronic low back pain (over a 2-year time period).

Grade C

With respect to the combination of anterior lumbar interbody fusion (ALIF) plus posterolateral fusion, there

Abbreviations used in this paper: ALIF = anterior lumbar interbody fusion; CCR = cost-to-charge ratio; CMS = Center for Medicare and Medicaid Services; CPT = Current Procedural Terminology; DRG = Diagnosis-Related Group; FRA = femoral ring allograft; HR-QOL = health-related quality of life; ICBG = iliac crest bone graft; ICD = International Classification of Diseases; ICER = incremental cost-effectiveness ratio; LOS = length of hospital stay; MIS = minimally invasive surgery; ODI = Oswestry Disability Index; QALY = quality-adjusted life year; rhBMP-2 = recombinant human bone morphogenetic protein–2; SEK = Swedish kronor; SF-36 = 36-Item Short Form Health Survey; SPORT = Spine Patient Outcomes Research Trial; TC = titanium cage; TDR = total lumbar disc replacement; TLIF = transforaminal lumbar interbody fusion; UK = United Kingdom; VAS = visual analog scale.

is Level II evidence that the use of a femoral ring allograft for interbody fusion is a more cost-effective interbody option than the use of a specific titanium cage.

From an economic perspective, both iliac crest bone graft and recombinant human bone morphogenetic protein–2 (rhBMP-2) are posterolateral fusion graft options in patients over the age of 60 (Level IV evidence).

From an economic perspective, both minimally invasive and open transforaminal lumbar interbody fusion (TLIF) techniques are options when treating patients with symptomatic Grade I degenerative spondylolisthesis (Level IV evidence).

Grade I

There are conflicting data regarding the cost-effectiveness of cell-salvage auto-transfusion as an adjunct to lumbar fusion (Level IV evidence).

Rationale

Ongoing changes in national health care policy have created an increased awareness on medical resource allocation and greater emphasis on cost-benefit analyses. As part of the American Recovery and Reinvestment Act of 2009, the federal government has allocated $1.1 billion in funds toward comparative-effectiveness research.[9,23] An area of specific interest has been the application of lumbar fusion in the management of degenerative spine disease, with a focus on establishing clinical efficacy and cost-effectiveness.[31] Management of chronic degenerative spinal conditions in the United States is estimated to cost nearly $85 billion annually, with a significant percentage attributed to the dramatic increase in the frequency of lumbar fusion procedures.[13,15,25] In 2004, more than 300,000 spinal fusions were performed in the US, accounting for more than $16 billion in hospital charges alone.[14] Advances in surgical fusion technologies have improved the surgeon's ability to attain a solid arthrodesis and expand the treatment options available for patients with spine disorders. Since the initial publication of the Lumbar Fusion Guidelines, there is recognition that the evolution of devices and techniques for lumbar fusion impacts not only surgical outcomes but also health care costs. The purpose of this qualitative review is to evaluate current research that examines the economic impact of lumbar fusion on the management of degenerative lumbar spine disease. The expense of fusion surgery and new fusion technologies must be weighed against the incremental improvement in patient outcomes and quality of life.

Search Criteria

A search of the National Library of Medicine database of literature was performed with limits: (("2002" [PDAT]: "2011"[PDAT]) AND English[lang]). Using the following terms: ("lumbar" AND "fusion"[All Fields]) yielded 4002 citations. The following terms were combined: ("lumbar" AND "fusion" AND "outcomes"[All Fields]), which yielded 807 citations and ("lumbar" AND "fusion" AND "cost"[All Fields]) which yielded 154 citations. The titles and abstracts of the 154 articles were reviewed. In addition, additional searches were performed with terms: (("lumbar fusion"[MeSH]) AND ("cost effectiveness"[MeSH]) OR ("employment status"[MeSH]), ("mortality") OR ("medical care costs[MeSH])" OR ("cost containment"[MeSH] OR "cost comparison"[MeSH]) OR ("spondylolisthesis")). Of the articles reviewed, 13 clinical series focusing on adult patients who underwent lumbar fusion procedures were selected based on the inclusion of an economic analysis.

A comprehensive economic analysis from a societal perspective that included multivariate sensitivity analyses was performed in 4 articles.[16,18,36,39] A cost analysis investigating various surgical approaches for lumbar fusion was performed in 1 study.[36] Another study performed a cost analysis on the type of interbody device used in lumbar fusion.[16] Two randomized trials compared outcomes and cost of lumbar fusion to conservative management.[19,39] Two studies examined the comparative cost-effectiveness of minimally invasive versus open TLIF.[43,44] One preliminary study provided cost-effectiveness data for TLIF procedures.[1] Six studies addressed incremental cost-effectiveness of new technology for lumbar fusion.[2,7,8,20,30,34]

Scientific Foundation

A cost-utility analysis is a specific type of cost-effectiveness evaluation that allows a comparison of 2 alternative treatment strategies in terms of the cost required for a given clinical outcome. These analyses are measured in terms of quality-adjusted life years (QALYs) gained, taking into account both the quantity and quality of life resulting from a given intervention.[21,32] For calculation of QALYs, patients must be surveyed using a preference-based health-related quality-of-life (HR-QOL) outcome instrument, such as the EQ-5D (EuroQol Group).[29,37] Another commonly used preference-based HR-QOL instrument is the SF-6D,[4] which consists of 11 items selected from the 36-Item Short Form Health Survey (SF-36).[5] An HR-QOL score is converted to a "health utility," typically a number on a continuum between 0, indicating death, and 1, indicating perfect health. Negative values can be generated when conditions considered worse than death exist.[38] A QALY is determined by the number of years in a given health state multiplied by the utility score assigned to that particular health state. A single year spent in perfect health is given the value of 1 QALY.

When comparing 2 treatment strategies, A and B, it is necessary to know the incremental cost-utility ratio (similar to the incremental cost-effectiveness ratio [ICER]). The ICER of Treatment B versus A is calculated as: (Cost of B – Cost of A)/(QALYs gained from B – QALYs gained from A).

From this calculation, the incremental cost of each additional QALY is determined when Treatment B is chosen over Treatment A. The acceptable cost per additional QALY represents society's willingness to pay and serves as a foundation for cost-effectiveness analyses. Since 1982, $50,000 per QALY gained has been cited as the threshold for a cost-effective intervention,[42] although

more recent proposals argue for a cut-off value closer to $100,000 or more, reflecting inflation and increased costs for research and development.[3]

There are 3 main categories of cost: direct, indirect, and intangible. Direct costs are resources that are consumed by the surgical procedure (i.e., operating room supplies, surgeon time and labor, cost of hospital stay). Indirect costs generally refer to a loss of productivity due to the morbidity or mortality of the surgical procedure. For example, the amount of work missed by the patient and/or their caretaker during the recovery period would qualify as indirect costs. Intangible costs include the pain and suffering from the surgical procedure. Both the indirect and the intangible costs are often difficult to quantify in monetary terms. Therefore, the total cost of a surgical procedure is based on the quantity of resources used and the assignment of cost to these resources. Determining the quantity of resources used is relatively straightforward for a surgical procedure. However, unit costs may vary between different countries, geographic regions, time periods, or hospitals. The cost perspective ("costs to whom") must be considered and expressly stated in any economic analysis. Costs to the patient for an intervention may be quite different from those to a hospital, a third-party payer, or to society itself.

Several methods have been introduced to estimate the total cost of a surgical procedure. One method utilizes the total hospital charge for the procedure and admission. The hospital charge is based on several data coding systems that are currently used to determine reimbursement, including the diagnosis-related group (DRG), International Classification of Diseases (ICD) system, and current procedural terminology (CPT) system.[10,12,33,45] This method, however, fails to reflect the actual amount of reimbursement received by the hospital or physician, or the actual costs, counting instead upon charges as a surrogate for costs. For the hospital charges of different centers to be used as a proxy for direct costs, a cost-to-charge ratio (CCR) must be calculated and applied. The CCR is specific for every hospital, for many departments within the hospital, and for a given time period. The CCR is calculated from Medicare Hospital Cost Reports (Worksheet C or D) in combination with claims data. Although obtaining CCRs is labor intensive, it has been suggested that this approach is the most accurate way to determine actual "cost" when comparing different centers.[17] Other methods for calculating cost of a surgical procedure include using total Medicare charges allowed, or the Medicare reimbursement. The most common and simplest way to estimate the direct cost for a procedure is using Medicare payments.[35,41] Real hospital costs can also be estimated by using the Center for Medicare and Medicaid Services (CMS) reimbursement value for DRG and CPT codes.

Micro-costing methodology involves measuring all the costs and benefits of a treatment as accurately as possible. It becomes particularly useful when evaluating and comparing regional differences in the resource utilization for a particular surgical procedure. However, micro-costing analysis is expensive and time-consuming due to extensive record keeping and database management. Often details, specific to a single institution, limit the generalizability of the conclusions. Finally, there are some elements of the micro-costing analysis that inevitably require estimation, which will also compromise the validity of the conclusions.[28]

An understanding of the methodology used to determine cost is critical when interpreting the results and conclusions of a study.[11] For example, in the recent Spine Patient Outcomes Research Trial (SPORT), the cost-effectiveness of surgery relative to nonoperative treatment for lumbar disc disease was $69,403/QALY using overall adult surgery costs (all payers), but only $34,355/QALY using Medicare population–specific surgery costs.[40] Using Medicare-based reimbursements will significantly lower the estimate of medical costs for any given treatment, which may be appropriate in older aged individuals, but irrelevant when considering younger patients with better (e.g., Workers Compensation) or worse (e.g., Medicaid) reimbursements.

Literature Review

Comparison of Lumbar Fusion to Nonoperative Therapy

Utilizing 2-year follow-up data from the SPORT study, Tosteson et al. investigated the cost-effectiveness of lumbar fusion for patients with degenerative spondylolisthesis and spinal stenosis.[39] Patients underwent nonoperative treatment, decompressive laminectomy, or laminectomy with fusion, with or without instrumentation and/or iliac crest bone graft. QALYs were calculated from EQ-5D scores at baseline, 6 weeks, and 3, 6, 12, and 24 months following treatment. Direct and indirect costs were collected prospectively based on 2004 Medicare payments. Operative management of spinal stenosis improved health significantly compared with nonoperative care, 0.17 QALYs gained, at a cost of $77,600 per QALY. Operative management of degenerative spondylolisthesis, 93% of which were lumbar fusions, provided significant benefit, with 0.23 QALYs gained at a cost of $115,600 per QALY. Although $115,600 is greater than the accepted societal expense per QALY, it is much less than previous estimates. Kuntz et al. reported that an instrumented lumbar fusion procedure cost $3,112,800 per QALY and instead favored noninstrumented fusion with a comparative medical benefit.[24] Although Tosteson et al. performed a rigorous cost analysis, the underlying heterogeneity of the study population and surgical techniques limits the validity of the study conclusions. In addition, follow-up data beyond 2 years will be essential to order to formulate meaningful recommendations regarding the cost-effectiveness of lumbar fusion over nonfusion treatments for lumbar degenerative disorders. This study provides Level II evidence that surgery for degenerative spondylolisthesis is effective but more costly than surgery for spinal stenosis (see Table 1).

Fritzell et al. performed a randomized controlled trial of patients with chronic low-back pain who underwent either lumbar fusion or nonoperative treatment.[19] Two hundred eighty-four patients from multiple centers were randomized, and outcomes were measured by the

TABLE 1: Economic outcome: summary of evidence*

Authors & Year	Level of Evidence	Description of Study	Conclusions
Soegaard et al., 2007	I	146 pts randomized to either PLF or circumferential fusion w/ long-term (4–8 yrs) EQ-5D utility outcomes assessment. Follow-up was 86%. Multi-way sensitivity analyses were performed. Circumferential lumbar fusion demonstrated clinical superiority over PLF. For each QALY gained by circumferential fusion, there was an incremental saving of $49,306.	Circumferential fusion was dominant over PLF; it is cost-effective.
Fritzell et al., 2011	I	152 pts w/ chronic LBP randomized to total disc replacement (TDR) or lumbar fusion. EQ-5D scores & costs were assessed at 2 yrs. Follow-up was 99%. Appropriate sensitivity analyses were performed. Outcomes & cost-effectiveness profiles were comparable.	Cost-effectiveness of TDR & lumbar fusion are comparable for chronic LBP.
Freeman et al., 2007	II	83 pts undergoing circumferential lumbar fusion, randomized to receive either titanium cages (TC) or femoral ring allograft (FRA). SF-6D used for cost-utility assessment. Follow-up was 94%. Multi-way sensitivity analyses were performed. At 2 yrs, mean QALY gain per pt was significantly greater for FRA (0.1914) than TC (0.0522). FRA was £1950 less expensive than TC.	The study was downgraded to Level II evidence because the study population was heterogeneous. FRA was dominant over a specific TC that was studied. FRA was cheaper & generated greater QALY gains.
Fritzell et al., 2004	III	284 pts w/ chronic LBP randomized to lumbar fusion or nonoperative treatment. Outcomes & costs were assessed at 2 yrs. 97% follow-up. Appropriate sensitivity analyses. QALYs gained were not calculated. Incremental cost of surgery for unit effect gained was SEK 2600 for global improvement; SEK 5200 for back pain by VAS; SEK 11,300 for ODI; SEK 4100 for return to work.	Outcomes were improved w/ fusion surgery compared to nonoperative treatment. Fusion surgery was more expensive than nonoperative treatment.
Savvidou et al., 2009	IV	50 pts undergoing PLF randomized to receive cell salvage autotransfusion (Group A) or allogenic transfusion (Group B). The cost of blood transfusion in Group A was (€995 ± €447) per pt & (€1220 ± €269) in Group B ($p < 0.05$).	The use of cell salvage autotransfusion results in less blood transfusion in elective lumbar fusions.
Reitman et al., 2004	IV	Retrospective cohort study of 102 pts who underwent lumbar fusion w/ either cell salvage autotransfusion or w/out autotransfusion (control). 36% of pts in the cell salvage group received additional allogenic transfusion compared to 50% of pts in control. Average blood-related cost was $512 for cell salvage compared to $270 in control.	Although cell salvage does decrease postoperative transfusions, it is more costly.
Carreon et al., 2009	IV	102 pts >60 yrs old requiring PLF were randomized to receive either rhBMP-2 or ICBG. Outcomes & costs were assessed at 2 yrs. SF-6D was used for utility measurements. Follow-up was 96%. No sensitivity analysis. The cost of fusion using rhBMP-2 was $39,967. Fusion w/ ICBG cost $42,286. In the ICBG group, there were 5 revision surgeries; in the rhBMP-2 cohort, 1 revision surgery. The 2 cohorts had similar improvement in SF-6D scores.	The cost of using rhBMP-2 is comparable to the cost of autograft for lumbar fusion surgery pts.
Glassman et al., 2008	IV	106 pts >60 yrs old randomized to either rhBMP-2 vs ICBG for anterior 1-level lumbar spine fusion. Outcomes & costs were assessed at 2 yrs. Total cost of care over 2 yrs was not significantly different ($42,574 for ICBG & $40,131 for rhBMP-2).	Use of rhBMP-2 is as effective as ICBG for lumbar spinal fusion w/ similar costs.
Alt et al., 2009	IV	Using pooled clinical data from previous study of 679 pts receiving either rhBMP-2 or autogenous bone graft for 1-level interbody lumbar fusion, the financial impact of rhBMP-2 was calculated from a societal perspective for 3 countries: Germany, France, & UK. Using a societal perspective (return- to work, secondary treatments costs, & reduced surgical time) rhBMP-2 used in ALIF surgery might reduce societal costs by €8483 in Germany, €9191 in France, & €8783 in UK.	The use of rhBMP-2 is associated w/ a reduction in return-to-work time & therefore appears to increase productivity from a societal perspective.

(continued)

TABLE 1: Economic outcome: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Conclusions
Cahill et al., 2009	IV	Retrospective cohort study of 328,468 pts from NIS who underwent spinal fusion, to determine complications & charges associated w/ use of rh-BMP. For lumbar fusion cases, LOS was 3.17% increased & total hospital charges were 31.30% greater following utilization of rh-BMP ($74,254 vs $57,393).	Rh-BMP is associated w/ increased LOS & total hospital charges in all categories of spinal fusion.
Adogwa et al., 2011	IV	45 pts who underwent a lumbar TLIF for spondylolisthesis (single-institution registry) were retrospectively reviewed. Lumbar TLIF was associated w/ a mean 2-yr cost of $42,854/QALY. No control group. No sensitivity analysis.	Preliminary economic data suggests lumbar TLIF might be cost-effective (<$50,000/QALY) at a 2-yr time horizon.
Wang et al., 2010	IV	Retrospective study of 74 pts treated w/ either minimally invasive or open TLIF lumbar procedures. Mean hospital charges for open surgery were higher than for minimally invasive TLIF. For 1-level fusion: MIS, $70,159 vs open TLIF, $78,444.	Open lumbar TLIF is associated w/ higher hospital charges than minimally invasive TLIF.
Wang et al., 2012	IV	Retrospective cohort study of 6106 pts from NIS who underwent MIS vs open TLIF. Hospital costs were calculated w/ CCRs. Single-level MIS & open TLIF were associated w/ comparable hospital costs. For 2-level TLIF: MIS ($33,879) was less costly than open surgery ($35,984).	Single-level minimally invasive TLIF & open TLIF have comparable hospital costs. For 2-level TLIF, MIS might be less costly than open surgery.

* ALIF = anterior lumbar interbody fusion; CCR = cost-to-charge ratio; FRA = femoral ring allograft; ICBG = iliac crest bone graft; LBP = low-back pain; LOS = length of stay; MIS = minimally invasive surgery; NIS = Nationwide Inpatient Sample; ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; pt = patient; QALY = quality-adjusted life year; rhBMP-2 = recombinant human bone morphogenetic protein–2; SEK = Swedish kronor; SF-6D = Short-Form 6D; SPORT = Spine Patient Outcomes Research Trial; TC = titanium cages; TLIF = transforaminal lumbar interbody fusion; UK = United Kingdom; VAS = visual analog scale.

Oswestry Disability Index (ODI) and visual analog scale (VAS) for 2 years. Patients who underwent surgery had 1 of 3 procedures: noninstrumented posterolateral fusion, instrumented posterolateral fusion, or a posterolateral circumferential fusion with pedicle screws and interbody grafts. Both direct and indirect costs were collected for each treatment group. The mean cost per patient was significantly higher in the surgical group (cost in Swedish kronor [SEK]: SEK 123,000 [US$18,731]) compared with the nonoperative group (SEK 65,200 [US$9929]). A significantly greater percentage of patients in the surgically treated group returned to part-time or full-time work compared with the nonoperative group (33% vs 16%, $p = 0.015$). Overall, lumbar fusion was associated with significantly greater improvements in pain and function compared with nonoperative treatment at 2 years. However, there was no difference in clinical outcome between the 3 fusion techniques studied. Compared with noninstrumented posterolateral fusion, the placement of pedicle screws increased hospital costs by 66% and a circumferential fusion procedure increased them by 103%. This study, however, did not measure QALYs, and therefore, it is difficult to interpret and generalize the data against more meaningful cost analyses.[26] Another significant limitation of this study is that the retrospective nature of the cost analysis, introducing the potential for significant recall bias regarding the patient's recollection of outpatient health resource utilization. Due to these design limitations, no definite conclusions regarding cost-effectiveness can be formulated from this economic analysis. It provides Level III evidence that outcomes are improved with fusion surgery for low-back pain compared with nonoperative treatment (see Table 1).

Lumbar Disc Arthroplasty Versus Posterior Lumbar Fusion

Fritzell's group performed a randomized controlled trial comparing the cost-effectiveness of total lumbar disc replacement (TDR) versus posterolateral fusion with or without interbody fusion.[18] One hundred fifty-two patients with chronic low-back pain were randomized to receive either TDR (80 patients) or fusion (72 patients). QALY outcomes were assessed using EQ-5D over a 2-year time period. Direct and indirect health costs were collected. Utilization of all outpatient resources and loss of work productivity were included to calculate health costs from a societal perspective. Multi-way sensitivity analyses were performed on excluding reoperations in both groups, costs for inpatient rehabilitation, and health cost discounting (see Table 1).

Follow-up data were available on 99% of cases at 2 years. From a societal perspective the mean health cost for TDR was SEK 599,560 (US$90,162) and for lumbar fusion was SEK 685,919 (US$103,149). This difference in cost was not statistically significant. At 2 years after surgery both groups demonstrated an improvement of 0.4 QALYs. The study's authors concluded that lumbar TDR and lumbar fusion have similar cost-effectiveness profiles in Sweden, although TDR was associated with lower costs from a health care perspective because of a lower reoperation rate at 2 years (see Table 1).

Circumferential Lumbar Fusion Versus Posterolateral Lumbar Fusion

Soegaard et al. randomized 146 patients with chronic low-back and leg pain to either posterolateral or circumferentially instrumented fusion and followed outcomes for 4–8 years after surgery.[36] Outcomes were measured with EQ-5D, SF-36, ODI, and pain scores. Service utilization (i.e., surgery, reoperations, rehospitalizations, general practitioner visits, etc.) was recorded per patient by the National Patient Registry in Denmark. Service utilization is valued by national average unit costs through the DRG system of coding. A micro-costing analysis was performed, including patient costs, medications, and productivity costs. The circumferentially treated group demonstrated better functional outcome ($p = 0.004$), higher fusion rate ($p < 0.04$), and fewer reoperations (15% versus 38%) compared with the posterolateral cohort. From a societal perspective, the circumferentially treated group demonstrated significantly lower costs compared with the posterolateral group ($p = 0.012$), primarily due to the higher reoperation rate and lower return to work rate observed in the posterolateral group. This study showed an incremental savings of $49,306 per QALY following a circumferential fusion compared with a posterolateral fusion. The study benefited from the long follow-up interval, large number of patients enrolled, and the comprehensive National Patient Registry. However, there are limitations, which include the heterogeneous population of patients with respect to presenting diagnosis and history of previous spine surgery. For example, patients with isthmic spondylolisthesis may have more favorable outcomes following lumbar fusion than other patients undergoing a lumbar fusion for different reasons. This study provides Level I evidence that circumferential fusion through a posterolateral approach is more cost-effective than stand-alone posterolateral fusion for up to 8 years following surgery (see Table 1).

Minimally Invasive Versus Open TLIF

Two retrospective studies by Wang et al. have addressed the cost-effectiveness of minimally invasive versus open TLIF.[43,44] In a retrospective review of 59 single-level TLIF cases at one institution using hospital charges as a surrogate for hospitalization costs, Wang and colleagues found that minimally invasive single-level TLIF was associated with lower hospital charges ($70,159) compared with open single-level TLIF ($78,444) ($p = 0.027$). Using the Nationwide Inpatient Sample (6106 cases) and applying CCRs to estimate hospital costs, Wang et al. found that minimally invasive single-level fusion was associated with hospital costs similar to those for open TLIF; however, for 2-level procedures, hospital costs of minimally invasive TLIF procedures ($33,879) were lower than costs of open surgery ($35,984; $p = 0.0023$). Neither study included outcome measures, so ICERs could not be calculated to assess cost-effectiveness (see Table 1).

A recent preliminary study reported cost-effectiveness data for TLIF using EQ-5D to measure QALYs gained over a 2-year time period. Adogwa et al. calculated health costs from a societal perspective by including inpatient and outpatient Medicare costs as well as workday losses.[1] While this type of analysis was limited because there was no comparison group, the results demonstrated that TLIF might be cost-effective ($42,854/QALY). Comparative studies are needed to calculate ICERs to validate these findings (see Table 1).

Type of Interbody Device Used for Lumbar Interbody Fusion

Freeman et al. randomized 83 patients with chronic low-back pain undergoing circumferential fusion to receive either a titanium cage (TC) or femoral ring allograft (FRA) in an effort to determine which interbody graft was more cost-effective.[16] Outcomes were measured by ODI and SF-6D at various intervals for up to 2 years following surgery. Direct cost data were derived from the National Health Service, using local center-specific unit costs for individual health resources. The indirect costs were measured by the human capital approach, which is based on the total expected production losses for an individual worker for the duration of disability.[22] The cost of TC was approximately 10 times higher than FRA (£1609.76 [US$2583] vs £158.92 [US$255]), and the mean total cost per patient was significantly higher in the TC group, £9052 (US$14,531) compared with £7102 (US$11,399) ($p < 0.001$). In addition, the mean QALY gained per patient over 24 months was significantly greater in the FRA group (0.1914) compared with the TC group (0.0522). Because the FRA proved to be less expensive and increased QALYs compared with a TC, an ICER was not necessary. Finally, using the human capital approach to cost productivity, total gross mean earnings in the 2 postoperative years were £7456 (US$11,968) in the TC group and £14,517 (US$23,303) in the FRA group. A higher percentage of FRA patients had a positive change in employment status following surgery compared with TC patients (21.6% vs 9.8%), and FRA patients were more likely to see an improvement in their salary after surgery. The strength of this study lies in its randomized design, excellent outcomes data, and comprehensive cost analysis. However, some utility data were missing, requiring estimation in the analysis. In addition, a few important health care resources were omitted from the cost analysis, including radiology costs, medication, outpatient services, and primary care costs. The clinical observations of Freeman et al. are supported by an earlier randomized trial by McKenna et al. that found improved clinical outcomes in patients implanted with FRA compared with TC.[27] These authors speculated that the improved pain relief with FRA may be due to the more physiological transfer of loads as the fusion matures compared with point loading with titanium cages. Overall, the study of Freeman et al. provides Level II evidence that the use of femoral ring allograft for lumbar interbody fusion is more cost-effective than the use of a specific titanium cage when performing a posterolateral circumferential fusion (see Table 1).

Cost-Effectiveness of rhBMP

Cahill et al. reviewed a retrospective cohort of 328,468 patients undergoing lumbar fusion procedures

from 2002 to 2006, selected from the Nationwide Inpatient Sample database, in an effort to determine the cost associated with the application of recombinant human bone morphogenetic protein (rhBMP) as a fusion enhancer.[7] The usage of rhBMP has increased from 0.69% of all fusions in 2002 to 24.89% of all fusions in 2006. The primary outcome measures included the rate of rhBMP utilization, complications, length of stay, and associated hospital charges. There were no differences in the rates of overall complications based on the application of rhBMP. The use of rhBMP was associated with an extended hospitalization (3.17% increase). The utilization of rhBMP was associated with a 20% increase in total hospital charges compared with those undergoing fusion without rhBMP ($74,254 vs $57,393). This analysis, however, is limited by the lack of clinical outcome data, and since only total hospital charges were considered, with a wide range in charges identified between different institutions, $54,737–$102,663 for lumbar fusions supplemented with rhBMP and $39,660–$83,608 when rhBMP was not included. This study provides Level IV evidence that rhBMP is associated with increased length of hospital stay (LOS) and total hospital charges in all categories of spinal fusion (see Table 1).

Glassman et al. conducted a randomized trial of rhBMP-2 versus iliac crest bone graft (ICBG) in patients over 60 years of age undergoing either single-level or multilevel instrumented posterolateral lumbar fusion.[20] Outcomes were measured up to 2 years after surgery utilizing validated outcomes instruments, including the ODI, SF-36, and numerical pain scales. Direct costs were obtained from actual reimbursement to the hospital. Where actual payments could not be determined, a Medicare fee schedule was used to assign direct cost. No significant differences in any of the outcome measures were identified between the 2 treatment groups. However, the fusion rate at 2 years, as measured by CT and presence of bridging bone, was significantly higher in the rhBMP-2 group (86.3% vs 70.8%, $p = 0.030$). Including revision surgery and outpatient costs (i.e., postoperative rehabilitation), the total cost of care over 2 years was not significantly different for 2 groups ($42,574 for the ICBG group vs $40,131 for the rhBMP-2 group). This study provides Level IV evidence that either ICBG or rhBMP-2 may be considered as posterolateral fusion graft options in patients over the age of 60 (see Table 1).

Using the same data as Glassman et al., Carreon et al. performed a cost-utility analysis of rhBMP-2 versus ICBG.[8] Utility was estimated from the SF-6D.[4] As described above, costs were determined from actual reimbursement to hospitals and physicians participating in the trial and included reimbursement for inpatient and outpatient services, radiographic imaging, and medications. There was no significant difference in change in utility between the 2 groups at any time point. The total cost of using rhBMP-2 was $39,967, with a 0.11 mean improvement in SF-6D; for ICBG, the cost was $42,286, with a mean improvement of 0.10 in SF-6D. The authors failed to account for indirect costs such as lost wages or out-of-pocket expenses. This study provides Level IV evidence that the cost of using rhBMP-2 is comparable to the cost of autograft for patients undergoing lumbar fusion surgery (see Table 1).

Alt et al. performed a cost analysis on a previously studied cohort of patients who either received rhBMP-2 or ICBG during lumbar fusion.[2,6] The study population included 279 patients randomized to either rhBMP-2 or ICBG and 400 patients from a prospective nonrandomized cohort. The original trial demonstrated significant reduction in surgery time (reduction of 54 minutes) and blood loss (reduction of 66 ml) when rhBMP-2 was used instead of ICBG. Patients in the rhBMP-2 group required fewer revision operations and returned to work earlier. Finally, the fusion rate and clinical outcomes of patients in the rhBMP-2 group were significantly better than those in the ICBG group. The authors performed a retrospective cost analysis from data collected in 3 countries: Germany, France, and the United Kingdom (UK). Direct costs were estimated from the expense associated with the use of rhBMP-2, operating room time, and revision surgery. Indirect costs were estimated from the loss of productivity and the national average of gross wages. From a societal perspective, the overall savings associated with use of rhBMP-2 compared with ICBG for Germany, France, and the UK were €8483 (US$11,745), €9191 (US$12,726), and €8783 (US$12,161), respectively. This study is limited by the retrospective study design and failure to account for costs associated with out-of-pocket expenses, medications, or outpatient treatment. This study provides Level IV evidence that the use of rhBMP-2 is associated with a reduction in return-to-work time and may increase productivity from a societal perspective.

Conflicting Data Regarding the Cost-Effectiveness of Cell-Salvage Autotransfusion

Savvidou et al. randomized 50 patients undergoing instrumented lumbar fusion to a group that received cell-salvage autotransfusion and a group that did not.[34] The total amount of allogeneic and cell-salvage blood used per patient was recorded. The cost for each unit of allogeneic blood (€450 [US$623]) versus cell-salvage blood (€370 [US$512]) was determined from the Greek ministry of health. The total transfusion cost in the group that received cell-salvage autotransfusion was significantly lower than the group that did not (€995 [US$1377] versus €1220 [US$1689], $p < 0.05$). Because this study did not measure outcomes, it is impossible to perform an adequate cost-effectiveness analysis regarding the use of autotransfusion. This study provides Level IV evidence that the use of cell-salvage autotransfusion lowered the costs of transfusing blood for elective lumbar fusions (see Table 1).

Reitman et al. performed a retrospective analysis of patients who had undergone instrumented lumbar fusion with and without cell-salvage autotransfusion.[30] There was no significant difference in blood loss between the 2 groups, and 38% of the blood in the study group was returned as cell-salvage autotransfusion. The average charge of the cell-salvage was $512 per patient compared with $270 per patient in the control group. The authors concluded that the cost of blood replacement was higher in the cell-salvage autotransfusion group. As with other studies investigating the application of cell-salvage autotransfu-

sion, there are insufficient data to perform a meaningful cost-effectiveness analysis. This study provides Level IV evidence that use of cell-salvage autotransfusion is more costly than normal postoperative transfusion (see Table 1).

Summary

Lumbar fusion for certain degenerative spine disorders can be effective in improving clinical outcomes and long-term quality of life when compared with nonoperative therapy. Comprehensive economic analyses that include long-term clinical outcomes data and both direct and indirect costs will be necessary before any recommendations can be made regarding the cost-effectiveness of various methods of lumbar fusion. Given the significant impact of lumbar degenerative disease, it is essential from a societal perspective that these studies be conducted. Recent cost-analyses have demonstrated the long-term benefits of circumferential fusions over posterolateral fusions, FRA over TC when performing an interbody fusion, and that both rhBMP-2 and ICBG are associated with similar costs.

Key Issues for Future Investigation

As new technologies for lumbar fusion are introduced and studies are performed to assess their effectiveness, the inclusion of an economic analysis is essential. Appropriate long-term follow-up is important when designing trials, as the benefits of lumbar fusion, both clinically and economically, may be apparent several years following the operation. The major challenge for investigators is to determine the most reliable estimate of cost. Medicare reimbursement may underestimate real costs, and medical charges may grossly overestimate true costs. New methodology is needed to allow for meaningful long-term assessment of health cost, and it may be that the best "laboratory" for these studies is a well-run health care system that follows costs directly along with patient outcomes as a matter of course.

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., 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 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: Ghogawala. 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: Ghogawala. Study supervision: Kaiser.

References

1. Adogwa O, Parker SL, Davis BJ, Aaronson O, Devin C, Cheng JS, et al: Cost-effectiveness of transforaminal lumbar interbody fusion for Grade I degenerative spondylolisthesis. Clinical article. **J Neurosurg Spine 15:**138–143, 2011
2. Alt V, Chhabra A, Franke J, Cuche M, Schnettler R, Le Huec JC: An economic analysis of using rhBMP-2 for lumbar fusion in Germany, France and UK from a societal perspective. **Eur Spine J 18:**800–806, 2009
3. Braithwaite RS, Meltzer DO, King JT Jr, Leslie D, Roberts MS: What does the value of modern medicine say about the $50,000 per quality-adjusted life-year decision rule? **Med Care 46:**349–356, 2008
4. Brazier J, Roberts J, Deverill M: The estimation of a preference-based measure of health from the SF-36. **J Health Econ 21:**271–292, 2002
5. Brazier J, Usherwood T, Harper R, Thomas K: Deriving a preference-based single index from the UK SF-36 Health Survey. **J Clin Epidemiol 51:**1115–1128, 1998
6. Burkus JK, Heim SE, Gornet MF, Zdeblick TA: Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device. **J Spinal Disord Tech 16:**113–122, 2003
7. Cahill KS, Chi JH, Day A, Claus EB: Prevalence, complications, and hospital charges associated with use of bone-morphogenetic proteins in spinal fusion procedures. **JAMA 302:**58–66, 2009
8. Carreon LY, Glassman SD, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, et al: RhBMP-2 versus iliac crest bone graft for lumbar spine fusion in patients over 60 years of age: a cost-utility study. **Spine (Phila Pa 1976) 34:**238–243, 2009
9. Chalkidou K, Whicher D, Kary W, Tunis S: Comparative effectiveness research priorities: identifying critical gaps in evidence for clinical and health policy decision making. **Int J Technol Assess Health Care 25:**241–248, 2009
10. Chulis GS: Assessing Medicare's prospective payment system for hospitals. **Med Care Rev 48:**167–206, 1991
11. Chumney EC, Biddle AK, Simpson KN, Weinberger M, Magruder KM, Zelman WN: The effect of cost construction based on either DRG or ICD-9 codes or risk group stratification on the resulting cost-effectiveness ratios. **Pharmacoeconomics 22:**1209–1216, 2004
12. Chute CG, Cohn SP, Campbell KE, Oliver DE, Campbell JR: The content coverage of clinical classifications. **J Am Med Inform Assoc 3:**224–233, 1996
13. Davis H: Increasing rates of cervical and lumbar spine sur-

gery in the United States, 1979-1990. **Spine (Phila Pa 1976) 19:**1117–1124, 1994
14. Deyo RA: Back surgery—who needs it? **N Engl J Med 356:**2239–2243, 2007
15. Deyo RA, Cherkin D, Conrad D, Volinn E: Cost, controversy, crisis: low back pain and the health of the public. **Annu Rev Public Health 12:**141–156, 1991
16. Freeman BJ, Steele NA, Sach TH, Hegarty J, Soegaard R: ISSLS prize winner: cost-effectiveness of two forms of circumferential lumbar fusion: a prospective randomized controlled trial. **Spine (Phila Pa 1976) 32:**2891–2897, 2007
17. Friedman B, De La Mare J, Andrews R, McKenzie DH: Practical options for estimating cost of hospital inpatient stays. **J Health Care Finance 29:**1–13, 2002
18. Fritzell P, Berg S, Borgström F, Tullberg T, Tropp H: Cost effectiveness of disc prosthesis versus lumbar fusion in patients with chronic low back pain: randomized controlled trial with 2-year follow-up. **Eur Spine J 20:**1001–1011, 2011
19. Fritzell P, Hägg O, Jonsson D, Nordwall A, Swedish Lumbar Spine Study Group: Cost-effectiveness of lumbar fusion and nonsurgical treatment for chronic low back pain in the Swedish Lumbar Spine Study: a multicenter, randomized, controlled trial from the Swedish Lumbar Spine Study Group. **Spine (Phila Pa 1976) 29:**421–434, 2004
20. Glassman SD, Carreon LY, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, et al: RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. **Spine (Phila Pa 1976) 33:**2843–2849, 2008
21. Gold MR, Franks P, McCoy KI, Fryback DG: Toward consistency in cost-utility analyses: using national measures to create condition-specific values. **Med Care 36:**778–792, 1998
22. Korthals-de Bos I, van Tulder M, van Dieten H, Bouter L: Economic evaluations and randomized trials in spinal disorders: principles and methods. **Spine (Phila Pa 1976) 29:**442–448, 2004
23. Kuehn BM: Institute of Medicine outlines priorities for comparative effectiveness research. **JAMA 302:**936–937, 2009
24. Kuntz KM, Snider RK, Weinstein JN, Pope MH, Katz JN: Cost-effectiveness of fusion with and without instrumentation for patients with degenerative spondylolisthesis and spinal stenosis. **Spine (Phila Pa 1976) 25:**1132–1139, 2000
25. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, et al: Expenditures and health status among adults with back and neck problems. **JAMA 299:**656–664, 2008
26. McCabe C, Claxton K, Culyer AJ: The NICE cost-effectiveness threshold: what it is and what that means. **Pharmacoeconomics 26:**733–744, 2008
27. McKenna PJ, Freeman BJ, Mulholland RC, Grevitt MP, Webb JK, Mehdian SH: A prospective, randomised controlled trial of femoral ring allograft versus a titanium cage in circumferential lumbar spinal fusion with minimum 2-year clinical results. **Eur Spine J 14:**727–737, 2005
28. Polsky D, Glick H: Costing and cost analysis in randomized controlled trials: caveat emptor. **Pharmacoeconomics 27:**179–188, 2009
29. Rabin R, de Charro F: EQ-5D: a measure of health status from the EuroQol Group. **Ann Med 33:**337–343, 2001
30. Reitman CA, Watters WC III, Sassard WR: The Cell Saver in adult lumbar fusion surgery: a cost-benefit outcomes study. **Spine (Phila Pa 1976) 29:**1580–1584, 2004
31. 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 3: assessment of economic outcome. **J Neurosurg Spine 2:**647–652, 2005
32. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC: The role of cost-effectiveness analysis in health and medicine. **JAMA 276:**1172–1177, 1996
33. Rutigliano MJ: Cost effectiveness analysis: a review. **Neurosurgery 37:**436–444, 1995
34. Savvidou C, Chatziioannou SN, Pilichou A, Pneumaticos SG: Efficacy and cost-effectiveness of cell saving blood autotransfusion in adult lumbar fusion. **Transfus Med 19:**202–206, 2009
35. Schulman KA: Medicare and cost-effectiveness analysis. **N Engl J Med 354:**207–209, 2006 (Letter)
36. Soegaard R, Bünger CE, Christiansen T, Høy K, Eiskjaer SP, Christensen FB: Circumferential fusion is dominant over posterolateral fusion in a long-term perspective: cost-utility evaluation of a randomized controlled trial in severe, chronic low back pain. **Spine (Phila Pa 1976) 32:**2405–2414, 2007
37. Suhonen R, Virtanen H, Heikkinen K, Johansson K, Kaljonen A, Leppänen T, et al: Health-related quality of life of day-case surgery patients: a pre/posttest survey using the EuroQoL-5D. **Qual Life Res 17:**169–177, 2008
38. Testa MA, Simonson DC: Assesment of quality-of-life outcomes. **N Engl J Med 334:**835–840, 1996
39. Tosteson AN, Lurie JD, Tosteson TD, Skinner JS, Herkowitz H, Albert T, et al: Surgical treatment of spinal stenosis with and without degenerative spondylolisthesis: cost-effectiveness after 2 years. **Ann Intern Med 149:**845–853, 2008
40. Tosteson AN, Skinner JS, Tosteson TD, Lurie JD, Andersson GB, Berven S, et al: The cost effectiveness of surgical versus nonoperative treatment for lumbar disc herniation over two years: evidence from the Spine Patient Outcomes Research Trial (SPORT). **Spine (Phila Pa 1976) 33:**2108–2115, 2008
41. Tumeh JW, Moore SG, Shapiro R, Flowers CR: Practical approach for using Medicare data to estimate costs for cost-effectiveness analysis. **Expert Rev Pharmacoecon Outcomes Res 5:**153–162, 2005
42. Ubel PA, Hirth RA, Chernew ME, Fendrick AM: What is the price of life and why doesn't it increase at the rate of inflation? **Arch Intern Med 163:**1637–1641, 2003
43. Wang MY, Cummock MD, Yu Y, Trivedi RA: An analysis of the differences in the acute hospitalization charges following minimally invasive versus open posterior lumbar interbody fusion. Clinical article. **J Neurosurg Spine 12:**694–699, 2010
44. Wang MY, Lerner J, Lesko J, McGirt MJ: Acute hospital costs after minimally invasive versus open lumbar interbody fusion: data from a US national database with 6106 patients. **J Spinal Disord Tech 25:**324–328, 2012
45. Yao P, Wiggs BR, Gregor C, Sigurnjak R, Dodek P: Discordance between physicians and coders in assignment of diagnoses. **Int J Qual Health Care 11:**147–153, 1999

Manuscript submitted March 11, 2014.
Accepted April 1, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14259.
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.

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 (κ = 0.62) than for facet fusion status (κ = 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% (κ = 0.25, $p < 0.0001$). Using majority consensus, 67% of cases were classified correctly as fused (93% sensitivity, 46% specificity). Anterior sentinel sign (κ = 0.34, $p < 0.0001$) showed 20% sensitivity & 92% specificity. Posterior sentinel sign (κ = 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.

J Neurosurg Spine 21:31–36, 2014

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.,[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] Praveen V. Mummaneni, M.D.,[1] Jeffrey C. Wang, M.D.,[10] and Michael G. Kaiser, M.D.[11]

[1]Department of Neurological Surgery, University of California, San Francisco, California; [2]Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, New York, New York; [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, 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

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

Abbreviations used in this paper: ALIF = anterior lumbar interbody fusion; DPQ = Dallas Pain Questionnaire; LBOS = Low Back Outcome Scale; LBPR = Low Back Pain Rating Scale; PLF = posterolateral lumbar fusion; VAS = visual analog scale.

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 isthmic 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 (pseudarthrosis). However, a number of authors have described patients with pseudarthrosis with favorable clinical outcomes and patients with solid osseous unions who have poor clinical outcomes.[3,7] 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.[4,8–10,14,18,19]

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 pseudarthrosis ($p = 0.01$), 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 pseudarthrosis can be generalized to patients with instrumented pseudarthrosis. 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,[21] Wetzel et al.,[20] and Djurasovic et al.[5] The remaining 2 studies—the study by Penta and Fraser[15] and the study by Epstein[6]—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*

Authors & Year	Level of Evidence	Description	Comment
Christensen et al., 2002	II	Prospective 2-yr follow-up of 148 pts randomized to PLF+PS or ALIF+PLF+PS. Clinical outcome was assessed w/ DPQ, LBPR, & work status survey scales. Both groups showed highly significant improvement in all 4 categories of life quality (DPQ), back pain, & 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.	Correlation was found btwn fusion status & functional outcome. Downgraded to Level III because of heterogeneous pt population (isthmic spondylolisthesis, DDD, previous surgeries) & use of static radiographs.
Kornblum et al., 2004	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 follow-up (81%). Standardized outcome measures & 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 ($p = 0.02$) & leg pain ($p = 0.001$).	Correlation found btwn fusion status & better pain scores. Study was graded Level II because it was retrospective.
Kim et al., 2006	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 & isthmic & degenerative spondylolisthesis. Validated outcome measures & dynamic radiographs. Of 17 pts w/ nonunion, 41% had superior clinical results as compared to 91% of pts w/ fusion.	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.
Thalgott et al., 2009	II	50 pts randomized to ALIF + post instrumentation w/ frozen vs freeze-dried FRA. 80% follow-up. Heterogeneous population: DDD, HNP, iatrogenic instability, spinal stenosis, degenerative spondylolisthesis. Study used dynamic radiographs, ODI, VAS. Freeze-dried allograft had significantly ($p = 0.026$) higher rate of pseudarthrosis. Differences in ODI, VAS btwn pts w/ & w/o fusion was not significant.	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.
Zdeblick, 1993	III	124 lumbar fusion pts were prospectively studied. Fusion status was determined using AP & flexion-extension radiography at 1 yr. Clinical results were rated as excellent, good, fair, or poor.	Groups w/ higher fusion rates did better. Downgraded to Level III due to heterogeneity of pt population & lack of validated outcomes.
Christensen et al., 1996	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, & 24% definitive pseudarthrosis. Pts w/ complete or questionable union had significantly better results than those w/ nonunion ($p < 0.01$).	DPQ scores correlated well w/ radiological outcome. Downgraded to Level III due to heterogeneity of pt population & use of static radiographs to assess fusion.
Penta & Fraser, 1997	III	103 ALIF pts (from a consecutive series of 125) had clinical (LBOS, VAS, MSPQ, ZDS) outcome assessment & 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 & was not influenced by the compensation status. Psychological disturbance at review & reop, 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 & becomes insignificant at 10 yrs.	Long-term (~10-yr) presence of radiological fusion was not associated w/ clinical outcome. Downgraded to Level III due to heterogeneity of pt population & retrospective nature of study.
Wetzel et al., 1999	III	74 consecutive cases of lumbar fusion. Nonvalidated outcome scores on pain relief & 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 & 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.	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 & use of nonvalidated outcome scale.

(continued)

TABLE 1: Correlation between radiographic outcome and function: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Comment
Epstein, 2008	III	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 & dynamic radiographs; 100% follow-up. Pseudarthrosis in 13 pts, leading to revision surgery in 1. Reports "nearly identical maximum improvement on SF-36" & no correlation btwn fusion & clinical outcome.	Found no correlation btwn pseudarthrosis & clinical results. The lesser quality prospective study was downgraded to Level III because of the heterogeneous pt population.
Djurasovic et al., 2011	III	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/ & w/o fusion; SF-36 & VAS (back & leg) scores did not show significant btwn-groups difference at 2 yrs. But 65% of pts w/ fusion achieved MCID on ODI vs 32% of pts w/o fusion ($p = 0.004$) & similar for SCB on ODI. 67% vs 50% reached SF-36 MCID ($p = 0.152$) & 63% vs 41% SCB for SF-36 ($p = 0.111$). Authors concluded that solid fusion contributes to better outcome.	Presence of solid fusion contributes to improved clinical outcome. Study was downgraded to Level III because of the heterogeneous pt population.

* 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.

eratively, fusion outcome was assessed using static plain radiography assessed by independent observers. These authors reported complete fusion in 52% of patients, questionable fusion in 24%, and definitive pseudarthrosis in 24%. Patients with complete or questionable union had significantly better DPQ scores than those with nonunion ($p < 0.01$). The authors concluded that DPQ scores correlated well with radiological outcome. This study is considered to provide Level III medical evidence supporting fusion status as a predictor of functional outcome because the radiographic and clinical follow-up evaluations were obtained at widely separated time points (between 3 and 11 years apart) and because the study relied on static plain radiography to determine fusion status and this modality has been shown to have limited accuracy.[16]

Wetzel and colleagues prospectively evaluated 74 consecutive patients who underwent lumbar fusion.[20] Outcomes were measured using subjective clinical outcome scores pertaining to pain relief and medication usage. The patients were observed at 5 intervals after surgery during a minimum 2-year follow-up period (range 24–35 months, mean 27 months). Fusion status was evaluated using lateral flexion-extension radiography in all cases, with the selective use of other techniques. The authors noted a 61% fusion rate. At final follow-up examination, 60% of patients had improvement in back pain and 70% had improvement in leg pain. The presence of radiographic fusion correlated positively with a successful clinical outcome ($r = 3.3$, $p = 0.010$). Similarly, in a prospective study of 124 lumbar fusion patients assigned to 3 different surgical treatment groups, Zdeblick assessed fusion status by performing static and flexion-extension lateral radiography at 1 year.[21] The clinical outcomes were rated as excellent, good, fair, or poor. The study showed that patients in the groups with higher fusion rates had better clinical outcomes. These studies, although prospective, are considered to provide Level III medical evidence in support of the correlation between radiographic and clinical outcome because of the use of nonvalidated clinical outcome measures.[19]

Djurasovic et al.[5] studied data on 193 patients collected from 3 clinical trials in which the patients underwent instrumented PLF for diverse indications. The authors compared outcomes in the patients with fusion versus those with nonunion and found that 65% of the patients with fusion achieved MCID on the ODI as compared with 32% of those with nonunion (a statistically significant difference).

In contrast, other studies have failed to demonstrate a statistically significant correlation between clinical and radiographic outcome in patients following lumbar arthrodesis surgery.

In a long-term outcome study (> 10 years), Penta and Fraser[15] reported on 103 patients who underwent ALIF (from a consecutive series of 125 cases). Clinical outcome assessment involved various validated outcome measures, including the Low Back Outcome Scale (LBOS). Eighty-seven patients also underwent fusion assessment with anteroposterior and lateral radiography. The authors reported that 78% of patients rated themselves as having "complete relief" or "a good deal of relief," but only 34%

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, Epstein[6] 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 instrument–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 Cockroft, 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 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: 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

1. Christensen FB, Hansen ES, Eiskjaer SP, Høy K, Helmig P, Neumann P, et al: Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubosset instrumentation: a prospective, randomized clinical study of 146 patients. **Spine (Phila Pa 1976) 27:**2674–2683, 2002
2. Christensen FB, Karlsmose B, Hansen ES, Bünger CE: Radiological and functional outcome after lumbar interbody spinal fusion. **Eur Spine J 5:**293–298, 1996
3. DePalma AF, Rothman RH: The nature of pseudoarthrosis. 1968. **Clin Orthop Relat Res (284):**3–9, 1992
4. Diedrich O, Perlick L, Schmitt O, Kraft CN: Radiographic characteristics on conventional radiographs after posterior lumbar interbody fusion: comparative study between radiotranslucent and radiopaque cages. **J Spinal Disord 14:**522–532, 2001
5. Djurasovic M, Glassman SD, Dimar JR II, Howard JM, Bratcher KR, Carreon LY: Does fusion status correlate with patient outcomes in lumbar spinal fusion? **Spine (Phila Pa 1976) 36:** 404–409, 2011
6. Epstein NE: An analysis of noninstrumented posterolateral lumbar fusions performed in predominantly geriatric patients using lamina autograft and beta tricalcium phosphate. **Spine J 8:**882–887, 2008
7. Flatley TJ, Derderian H: Closed loop instrumentation of the lumbar spine. **Clin Orthop Relat Res (196):**273–278, 1985
8. Gaetani P, Aimar E, Panella L, Levi D, Tancioni F, Di Ieva A, et al: Functional disability after instrumented stabilization in lumbar degenerative spondylolisthesis: a follow-up study. **Funct Neurol 21:**31–37, 2006
9. Greenough CG, Peterson MD, Hadlow S, Fraser RD: Instrumented posterolateral lumbar fusion. Results and comparison with anterior interbody fusion. **Spine (Phila Pa 1976) 23:**479–486, 1998
10. Hackenberg L, Halm H, Bullmann V, Vieth V, Schneider M, Liljenqvist U: Transforaminal lumbar interbody fusion: a safe technique with satisfactory three to five year results. **Eur Spine J 14:**551–558, 2005
11. Kaiser MG, Eck JC, Groff MW, Watters WC III, Dailey AT, Resnick DK, et al: Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: Introduction and methodology. **J Neurosurg Spine 21:** 2–6, 2014
12. Kim KT, Lee SH, Lee YH, Bae SC, Suk KS: Clinical outcomes of 3 fusion methods through the posterior approach in the lumbar spine. **Spine (Phila Pa 1976) 31:**1351–1358, 2006
13. Kornblum MB, Fischgrund JS, Herkowitz HN, Abraham DA, Berkower DL, Ditkoff JS: Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective long-term study comparing fusion and pseudarthrosis. **Spine (Phila Pa 1976) 29:**726–734, 2004
14. Lee CS, Hwang CJ, Lee SW, Ahn YJ, Kim YT, Lee DH, et al: Risk factors for adjacent segment disease after lumbar fusion. **Eur Spine J 18:**1637–1643, 2009

15. Penta M, Fraser RD: Anterior lumbar interbody fusion. A minimum 10-year follow-up. **Spine (Phila Pa 1976) 22:**2429–2434, 1997
16. 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 5: correlation between radiographic and functional outcome. **J Neurosurg Spine 2:**658–661, 2005
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
18. Tokuhashi Y, Matsuzaki H, Oda H, Uei H: Clinical course and significance of the clear zone around the pedicle screws in the lumbar degenerative disease. **Spine (Phila Pa 1976) 33:**903–908, 2008
19. Vamvanij V, Fredrickson BE, Thorpe JM, Stadnick ME, Yuan HA: Surgical treatment of internal disc disruption: an outcome study of four fusion techniques. **J Spinal Disord 11:**375–382, 1998
20. Wetzel FT, Brustein M, Phillips FM, Trott S: Hardware failure in an unconstrained lumbar pedicle screw system. A 2-year follow-up study. **Spine (Phila Pa 1976) 24:**1138–1143, 1999
21. Zdeblick TA: A prospective, randomized study of lumbar fusion. Preliminary results. **Spine (Phila Pa 1976) 18:**983–991, 1993

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.

J Neurosurg Spine 21:37–41, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 6: Discography for patient selection

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

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

Identifying the etiology of pain for patients suffering from chronic low-back pain remains problematic. Noninvasive imaging modalities, used in isolation, have not consistently provided sufficient evidence to support performance of a lumbar fusion. Provocative testing has been used as an adjunct in this assessment, either alone or in combination with other modalities, to enhance the diagnostic capabilities when evaluating patients with low-back pain. There have been a limited number of studies investigating this topic since the publication of the original guidelines. Based primarily on retrospective studies, discography, as a stand-alone test, is not recommended to formulate treatment strategies for patients with low-back pain. A single randomized cohort study demonstrated an improved potential of discoblock over discography as a predictor of success following lumbar fusion. It is therefore recommended that discoblock be considered as a diagnostic option. There is a possibility, based on a matched cohort study, that an association exists between progression of degenerative disc disease and the performance of a provocative discogram. It is therefore recommended that patients be counseled regarding this potential development prior to undergoing discography.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14269)

Key Words • **fusion** • **lumbar spine** • **discography** • **discoblock** • **practice guidelines**

Recommendations

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

Grade C

It is recommended that discoblock be considered as a diagnostic option during the evaluation of a patient presenting with chronic low-back pain (single Level II study).

It is recommended that lumbar discography not be used as a stand-alone test on which treatment decisions are based for patients with low-back pain with abnormal imaging studies (single Level II study).

It is recommended that within the discussion of potential risks for patients undergoing provocative discography, the potential for acceleration of the degenerative process be included as there is evidence to suggest an association between advanced degenerative spondylosis and a history of undergoing provocative discography.

Rationale

Surgical intervention for the treatment of chronic low-back pain has demonstrated inconsistent and less favorable results than procedures performed for other degenerative spine disorders. This is in part due to an inability to ac-

Abbreviations used in this paper: AAOS = American Academy of Orthopaedic Surgeons; JOA = Japanese Orthopaedic Association; MODEMS = Musculoskeletal Outcomes Data Evaluation and Management System; ODI = Oswestry Disability Index; VAS = visual analog scale.

curately determine the specific source of a patient's pain. It is well known that degenerative changes identified on an MRI may occur in asymptomatic patients, and therefore such findings cannot be used as the sole justification for surgery and are not predictive of clinical outcome.[1]

In an effort to isolate the source of pain, provocative testing (intended to reproduce a patient's pain), has been integrated into the evaluation of patients presenting with chronic low-back pain. Discography has been used in conjunction with MRI in an attempt to better identify specific patients who might benefit from lumbar fusion for chronic low-back pain. The purpose of this review is to examine the medical evidence regarding the utilization of discography as a diagnostic modality in the evaluation of patients presenting with chronic low-back pain being considered for lumbar spine fusion. Additional information regarding the methodologies and criteria used to evaluate the evidence discussed below is located in the first article in this issue (Part 1: Introduction and methodology).[12]

Literature Search

The database of the National Library of Medicine was searched for articles published between July 2003 and December 2011 using the following search terms: ((discography OR discogram) AND lumbar fusion AND (patient selection OR predictive value of tests) AND ((“2003”[PDat]: “3000”[PDat]) AND (Humans[MeSH]) AND (English[lang]))) OR ((discography OR discogram) AND (((“Lumbosacral Region”[MeSH] OR “Lumbar Vertebrae”[MeSH]) AND “Spinal Fusion”[MeSH]) OR “lumbar fusion”[All Fields] OR (“lumbar”[title] AND “fusion”[title])) AND (patient selection OR predictive value of tests) AND ((“2003”[PDat]: “3000”[PDat]) AND (Humans[MeSH]) AND (English[lang])))). Search results were limited to human studies, English language, and age between 18 and 65 years. The titles and abstracts of these articles were reviewed, and duplicates, technical notes, reviews, and papers that did not describe the use of discography for the diagnosis and management of patients with low-back pain were discarded. The reference lists of the remaining articles were inspected, and additional relevant papers were identified. From this group of citations, 6 were selected as the most relevant and are briefly described in the evidentiary table (see Table 1). The remaining references provided additional background information and are included in the bibliography.

Scientific Foundation

The use of discography for the diagnosis of lumbar intervertebral disc abnormalities in patients with low-back pain has been well described.[9,15] The key components of discography that aid in the diagnosis of patients with low-back pain include a reproduction of the patient's concordant pain, visualization of the disc morphology, and injection pressures. If each of these factors is found to suggest symptomatic disc degeneration, the test is considered to be positive. By recreating the patient's pain, proponents of discography argue that it is more sensitive and specific than other imaging modalities, including plain radiographs, myelography, and MRI, which are known to identify both symptomatic and asymptomatic abnormalities.[2–4] However, critics question the reliability and specificity of discography since concordant pain has been suggested to originate from nonspine sources and can be reproduced in patients without any prior history of back pain.[5,8,11]

A prospective study intended to evaluate the predictive value of provocative discography following lumbar fusion was performed by Carragee et al.[7] Lumbar fusions were performed in 32 patients with presumed discogenic pain and a positive discogram (see Table 1). A circumferential fusion was performed in a single day and consisted of an anterior lumbar interbody fusion with femoral ring allograft, buttress screw, and local bone/ allograft inserted within and around the structural graft. The posterior approach included an intertransverse fusion with allograft and pedicle screw instrumentation. Clinical outcomes were measured using a visual analog scale (VAS), American Academy of Orthopaedic Surgeons (AAOS) Musculoskeletal Outcomes Data Evaluation and Management System (MODEMS) instrument, Oswestry Disability Index (ODI), psychological testing, and the Fear Avoidance and Behavior Questionnaire. Clinical success at the 2-year follow-up point was defined by a VAS score of 2 or less, an ODI score of less than 15, return to work, no use of narcotic medications, and no daily pain medication requirement. Surgical comorbidity was compared with a control cohort of 34 patients undergoing lumbar fusion for unstable isthmic spondylolisthesis. Successful outcomes were observed in 72% of the control spondylolisthesis group as compared with only 27% in the discography group. The percentage of patients achieving a minimally acceptable improvement in the control group was 91% versus 43% for the discography group. The positive predictive value of discography was estimated to be between 50% and 60%. This study suffered from several design limitations, including the lack of an appropriate control group and is therefore considered to provide only Level IV evidence against utility of discography as a predictive tool.

Wetzel et al. performed a retrospective case review of 48 patients with a diagnosis of discogenic low-back pain based on positive discography and CT or MRI.[18] All patients underwent lumbar fusion of the levels determined to be symptomatic based on discography. The number of levels fused ranged from 1 to 4, and a wide variety of fusion techniques were used, including anterior or posterior approaches, with or without instrumentation. Clinical outcomes were subjectively graded as excellent, good, fair, or poor. Radiographic evidence of fusion was defined as 4° or less of motion on flexion-extension radiographs with the presence of mature trabecular bone across all levels. Twenty-three (47.9%) went on to successful radiographic fusion, and 22 of those had a satisfactory clinical outcome. Overall, 46% of patients were found to have a satisfactory clinical outcome. This study provided Level IV evidence against the use of discography in predicting clinical success following lumbar spine fusion. Limitations of the study included a variety of surgical techniques and lack of quantitative and validated clinical outcomes measures (see Table 1).

A retrospective case series of 53 patients undergoing

TABLE 1: Discography for patient selection: summary of evidence*

Authors & Year	Level of Evidence	Description	Comments
Ohtori et al., 2009	II	42 pts undergoing noninstrumented ALIF w/ iliac crest autograft were evaluated w/ discography or discoblock in a randomized fashion. 12 were eliminated due to a lack of response w/ either technique. VAS, ODI, & JOA scores were obtained preoperatively & at 1, 2, & 3 yrs. All pts achieved solid fusion by 2 yrs based on plain radiographs & CT. Significantly improved pain & disability scores were achieved in pts w/ a positive discoblock vs those w/ positive discography ($p < 0.05$). The authors concluded that discoblock was a superior predictor of success following lumbar fusion vs discography.	This study provided Level I evidence for the benefits of the discoblock technique over the traditional discogram technique. It is downgraded to Level II due to failure to perform a power analysis to determine sample size for randomization & the small sample size for each cohort.
Carragee et al., 2009	II	This prospective, matched-cohort study was performed to determine if provocative discography was associated w/ DDD progression. 75 pts undergoing provocative discography underwent baseline MRI & were compared to a matched cohort. At 10 yrs following the procedure, 50 discography pts & 52 controls underwent repeat MRI. All qualitative & quantitative parameters were more advanced in the discography group. The authors concluded that provocative discography was associated w/ progression of DDD & should be considered in discussion of risks.	This prognostic study is limited by significant loss to follow-up; over 30% of pts were not available for final MRI evaluation. It is therefore downgraded to Level II evidence that an association exists between provocative discography & progression of lumbar DDD.
Gill & Blumenthal, 1992	II	Retrospective comparative study of 53 pts who underwent ALIF at L5–S1. All had preop MRI & discography & follow-up of at least 24 mos. 19 had Type I discography findings of an anular tear that did not extend to the periphery of the disc w/ normal MRI. The remaining pts had Type II or III discography finding of extension of the anular tear to the periphery of the disc w/ abnormal MRI. Clinical outcomes were ODI, VAS, & pain drawing. Successful functional outcome was return to work or normal activities & no use of narcotics. In pts w/ Type II or III discography findings & abnormal MRI there was a 75% success vs 50% in those w/ Type I discography findings & normal MRI.	This study provided Level II evidence against use of discography in the presence of a normal MRI scan in predicting results following lumbar fusion.
Carragee et al., 2006	IV	Lumbar fusion performed in 32 pts w/ discogenic pain & positive discogram. Circumferential fusion performed w/ ALIF & FRA combined w/ intertransverse fusion w/ allograft & pedicle screw instrumentation. Clinical outcomes were measured w/ VAS, AAOS MODEMS instrument, ODI, psychological testing, & Fear Avoidance and Behavior Questionnaire. Clinical success at 2-yr follow-up was defined by VAS score ≤2, ODI score <15, return to work, no narcotics, & no daily pain medication requirements. Clinical results compared to results in control cohort of 34 pts undergoing lumbar fusion for spondylolisthesis. Successful outcomes observed in 72% of control group vs 27% of discography group. The percentage of pts achieving a minimally acceptable improvement in the control group was 91% vs 43% for the discography group. Positive discography had a positive predictive value of 50–60%.	This case series provides Level IV evidence against utility of discography as a predictive tool.
Wetzel et al., 1994	IV	Retrospective case series involving 48 pts w/ discogenic back pain w/ positive discography & CT or MRI. All had lumbar fusion of the levels determined based on discography. Number of levels fused ranged from 1 to 4, w/ a wide variety of techniques. Clinical outcomes were subjectively graded as excellent, good, fair, or poor. Radiographic fusion defined as 4° or less motion on flexion-extension radiographs & mature trabecular bone across all levels. 23 (47.9%) went on to successful radiographic fusion, & 22 of those had a satisfactory clinical outcome. Overall, 46% had satisfactory clinical outcome.	This study provided Level IV evidence against the use of discography in predicting clinical success following lumbar spine fusion. Limitations of the study included a variety of surgical techniques & lack of quantitative & validated clinical outcomes measures.
Knox & Chapman, 1993	IV	Retrospective case series of 22 ALIF pts w/ positive discogram. Subjective clinical outcomes were reported as good in 35% of pts, fair in 18%, & poor in 47%.	This study provided Level IV evidence against the use of discography in predicting clinical success following lumbar spine fusion. Limitations of this study included a small sample size & lack of quantitative & validated clinical outcomes measures.

* AAOS = American Academy of Orthopaedic Suregeons; ALIF = anterior lumbar interbody fusion; DDD = degenerative disc disease; FRA = femoral ring allograft; JOA = Japanese Orthopaedic Association; MODEMS = Musculoskeletal Outcomes Data Evaluation and Management System; ODI = Oswestry Disability Index; pts = patients; VAS = visual analog scale.

anterior lumbar interbody fusion at L5–S1 was performed by Gill and Blumenthal.[10] All patients had preoperative MRI and discography and at least 24 months of follow-up. Nineteen of the patients had Type I discography findings of an anular tear that did not extend to the periphery of the disc with normal MRI findings. The remaining patients had a Type II or III discography finding of extension of the anular tear to the periphery of the disc with abnormal MRI findings. Clinical outcomes were measured by the "Oswestry Pain Questionnaire" (ODI), VAS, and a pain drawing. A successful functional outcome was defined as the ability to return to work or normal activities and no use of narcotic medications. In the patients with Type II or III discography findings and abnormal MRI findings, there was a 75% success rate postoperatively, while there was a 50% success rate in those with Type I discography findings and normal MRI findings. This study provided Level IV evidence against the use of discography in the presence of normal MRI findings in predicting results following lumbar fusion. The limitations of this study include being a retrospective case series and having a small sample size (see Table 1).

Additional Level IV evidence against the use of discography for the prediction of clinical success following lumbar fusion was provided by Knox and Chapman.[14] They performed a retrospective analysis of a case series involving 22 patients who had positive discogram findings and underwent anterior lumbar interbody fusion. Subjective clinical outcomes were reported as good in 35% of patients, fair in 18%, and poor in 47%. Limitations of this study included a small sample size and lack of quantitative and validated clinical outcomes measures.

Willems et al. conducted a retrospective review of a series of cases to determine whether preoperative discography of adjacent-level discs could predict clinical outcome in patients undergoing lumbar fusion.[19] This study began with 209 patients, but 12 were eliminated for lack of data, and an additional 115 received conservative treatment. The remaining 82 patients had lumbar fusion and their cases were used in the analysis. Outcomes measures included a VAS pain scale and Odom's criteria. The preoperative discography results for adjacent levels did not affect clinical outcomes in this series of patients. This study provides Level IV evidence against the use of adjacent-level discography as a predictor of clinical success after lumbar fusion.

An alternative to the traditional technique of discography is the technique known as a "discoblock," which involves injecting the disc with an anesthetic agent instead of a contrast agent in an effort to eliminate as opposed to reproducing a patient's pain. This modified technique was compared with traditional discography by Ohtori et al.[16] Forty-two patients undergoing a noninstrumented anterior lumbar interbody fusion with iliac crest autograft were evaluated preoperatively with either provocative discography or discoblock in a randomized fashion. Twelve patients were eliminated from the study due to a lack of response with either technique. Outcome measures including VAS, ODI, and Japanese Orthopaedic Association (JOA) scores were obtained preoperatively and at 1, 2, and 3 years postoperatively. All patients reportedly achieved a solid fusion by 2 years following surgery based on plain radiographs and CT scans. Significantly improved clinical outcomes with respect to both pain and disability scores were achieved following lumbar fusion in patients with positive discoblock findings as compared with those with positive discography findings ($p < 0.05$). The authors suggested that the discoblock technique proved to be a better predictor of success following lumbar fusion than provocative discography. This study provided Level II evidence for the benefits of the discoblock technique over the traditional provocative discography. Limitations of the study included a small sample size and lack of a power analysis.

More recently, concern has developed over the possibility that diagnostic disc injections may lead to iatrogenic injury to the disc and accelerate the rate of disc degeneration. Animal studies have demonstrated degeneration of an intervertebral disc due to a needle puncture of the annulus fibrosis.[13,17] Carragee et al. conducted a comparative prospective cohort study over a 10-year period, during which they followed the progression of disc degeneration in patients with and without a history of discography.[6] MRI studies were performed at baseline and 10 years following provocative discography. Two blinded radiologists and 2 blinded orthopedic surgeons evaluated the images. All outcome measures, including progression of disc degeneration, occurrence of new herniations, loss of disc height, and loss of disc signal intensity, were found to be significantly worse in the patients who had undergone discography. Of the original sample cohort of 150 patients enrolled, only 68% were available at the time of follow-up. This study was classified as a prognostic study regarding the outcome after performance of a provocative discography and was downgraded to Level II evidence of the potential risk associated with the use of provocative discography due to the 32% loss to follow-up.

Summary

The use of discography to aid in patient selection for lumbar fusion remains controversial. Based on the current literature, there is insufficient evidence to suggest that discography should be used as an independent predictor of success following lumbar fusion for low-back pain. There is limited evidence to suggest that a discoblock or anesthetizing the disc instead of injecting contrast material provides superior predictive value. More recent evidence, however, suggests a possible risk of discography leading to an acceleration of disc degeneration.

Key Issues for Further Investigation

Determining the diagnostic potential of a specific test rests on the ability to compare results between a "gold standard" and the test under investigation. Due to the lack of a "gold standard" when attempting to identify the source of a patient's low-back pain, such evaluations are extremely difficult to design and interpret. With respect to discography, there remains insufficient evidence to support its routine use as a diagnostic modality when evaluating patients with low-back pain. While further investigation may help elucidate the potential of discog-

raphy, such trials may not be feasible given the recent evidence that accelerated rates of disc degeneration are associated with a history of previous discograms.

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., 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 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: Eck. 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: Eck. Study supervision: Kaiser.

References

1. Boden SD, McCowin PR, Davis DO, Dina TS, Mark AS, Wiesel S: Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. **J Bone Joint Surg Am 72:**1178–1184, 1990
2. Braithwaite I, White J, Saifuddin A, Renton P, Taylor BA: Vertebral end-plate (Modic) changes on lumbar spine MRI: correlation with pain reproduction at lumbar discography. **Eur Spine J 7:**363–368, 1998
3. Brightbill TC, Pile N, Eichelberger RP, Whitman M Jr: Normal magnetic resonance imaging and abnormal discography in lumbar disc disruption. **Spine (Phila Pa 1976) 19:**1075–1077, 1994
4. 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
5. Carragee EJ, Chen Y, Tanner CM, Truong T, Lau E, Brito JL: Provocative discography in patients after limited lumbar discectomy: a controlled, randomized study of pain response in symptomatic and asymptomatic subjects. **Spine (Phila Pa 1976) 25:**3065–3071, 2000
6. Carragee EJ, Don AS, Hurwitz EL, Cuellar JM, Carrino JA, Herzog R: 2009 ISSLS Prize Winner: Does discography cause accelerated progression of degeneration changes in the lumbar disc: a ten-year matched cohort study. **Spine (Phila Pa 1976) 34:**2338–2345, 2009 [Erratum in **Spine (Phila Pa 1976) 35:**1414, 2010]
7. Carragee EJ, Lincoln T, Parmar VS, Alamin T: A gold standard evaluation of the "discogenic pain" diagnosis as determined by provocative discography. **Spine (Phila Pa 1976) 31:**2115–2123, 2006
8. Carragee EJ, Tanner CM, Khurana S, Hayward C, Welsh J, Date E, et al: The rates of false-positive lumbar discography in select patients without low back symptoms. **Spine (Phila Pa 1976) 25:**1373–1381, 2000
9. Collis JS Jr, Gardner WJ: Lumbar discography. An analysis of one thousand cases. **J Neurosurg 19:**452–461, 1962
10. Gill K, Blumenthal SL: Functional results after anterior lumbar fusion at L5-S1 in patients with normal and abnormal MRI scans. **Spine (Phila Pa 1976) 17:**940–942, 1992
11. Holt EP Jr: The question of lumbar discography. **J Bone Joint Surg Am 50:**720–726, 1968
12. Kaiser MG, Eck JC, Groff MW, Watters III WC, Dailey AT, Resnick DK, et al: Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: Introduction and methodology. **J Neurosurg Spine 21:**2–6, 2014
13. Kim KS, Yoon ST, Li J, Park JS, Hutton WC: Disc degeneration in the rabbit: a biochemical and radiological comparison between four disc injury models. **Spine (Phila Pa 1976) 30:**33–37, 2005
14. Knox BD, Chapman TM: Anterior lumbar interbody fusion for discogram concordant pain. **J Spinal Disord 6:**242–244, 1993
15. Lindblom K: Technique and results of diagnostic disc puncture and injection (discography) in the lumbar region. **Acta Orthop Scand 20:**315–326, 1951
16. Ohtori S, Kinoshita T, Yamashita M, Inoue G, Yamauchi K, Koshi T, et al: Results of surgery for discogenic low back pain: a randomized study using discography versus discoblock for diagnosis. **Spine (Phila Pa 1976) 34:**1345–1348, 2009
17. Sobajima S, Kompel JF, Kim JS, Wallach CJ, Robertson DD, Vogt MT, et al: A slowly progressive and reproducible animal model of intervertebral disc degeneration characterized by MRI, X-ray, and histology. **Spine (Phila Pa 1976) 30:**15–24, 2005
18. Wetzel FT, LaRocca SH, Lowery GL, Aprill CN: The treatment of lumbar spinal pain syndromes diagnosed by discography. Lumbar arthrodesis. **Spine (Phila Pa 1976) 19:**792–800, 1994
19. Willems PC, Elmans L, Anderson PG, van der Schaaf DB, de Kleuver M: Provocative discography and lumbar fusion: is preoperative assessment of adjacent discs useful? **Spine (Phila Pa 1976) 32:**1094–1100, 2007

Manuscript submitted March 13, 2014.
Accepted April 1, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14269.
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.

J Neurosurg Spine 21:42–47, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 7: Lumbar fusion for intractable low-back pain without stenosis or spondylolisthesis

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

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

Establishing an appropriate treatment strategy for patients presenting with low-back pain, in the absence of stenosis or spondylolisthesis, remains a controversial subject. Inherent to this situation is often an inability to adequately identify the source of low-back pain to justify various treatment recommendations, such as lumbar fusion. The current evidence does not identify a single best treatment alternative for these patients. Based on a number of prospective, randomized trials, comparable outcomes, for patients presenting with 1- or 2-level degenerative disc disease, have been demonstrated following either lumbar fusion or a comprehensive rehabilitation program with a cognitive element. Limited access to such comprehensive rehabilitative programs may prove problematic when pursuing this alternative. For patients whose pain is refractory to conservative care, lumbar fusion is recommended. Limitations of these studies preclude the ability to present the most robust recommendation in support of lumbar fusion. A number of lesser-quality studies, primarily case series, also support the use of lumbar fusion in this patient population. *(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14270)*

Key Words • **low-back pain** • **lumbar fusion** • **lumbar spondylosis** • **practice guidelines** • **lumbar spine**

Recommendations

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

Grade B

Lumbar fusion or a comprehensive rehabilitation program incorporating cognitive therapy are recommended as treatment alternatives for patients with chronic low-back pain that is refractory to traditional conservative treatment, such as physical therapy, and is due to 1- or 2-level degenerative disc disease without stenosis or spondylolisthesis (multiple Level II studies).

It is recommended that lumbar fusion be performed for patients whose low-back pain is refractory to conservative treatment (physical therapy or other nonoperative measures) and is due to 1- or 2-level degenerative disc disease without stenosis or spondylolisthesis (multiple Level II studies).

Rationale

Lumbar fusion has become an accepted treatment alternative for low-back pain associated with stenosis and spondylolisthesis. There is a growing body of evidence including that from the Spine Patient Outcomes Research Trial (SPORT) that consistently demonstrates improved

Abbreviations used in this paper: ODI = Oswestry Disability Index; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale.

clinical outcomes with lumbar fusions for patients who fail conservative care.[1,32]

Chronic low-back pain associated with lumbar spondylosis, in the absence of stenosis or spondylolisthesis, is a common clinical problem; however, the optimal treatment strategy for this condition remains a controversial topic. Part of this uncertainty results, in many cases, from the inability to accurately determine the actual source of pain. The lack of specificity regarding the changes identified on MRI only adds to the uncertainty when formulating a management strategy. There are many conservative and several surgical treatment options available for the treatment of chronic back pain; however, consistent evidence of superior efficacy of one approach over another is lacking. When conservative measures fail to improve the patient's pain, lumbar fusion is often considered an appropriate treatment alternative. The high costs, risk of serious complications, and lack of consistent supporting evidence raise questions as to whether fusion for lumbar spondylosis is cost-effective and will lead to functional recovery. The purpose of this review is to evaluate the published literature regarding the use of lumbar fusion for the treatment of patients with intractable low-back pain without stenosis or spondylolisthesis.

Literature Search

The database of the National Library of Medicine was searched for articles published between July 2003 and December 2011 using the following search terms: ("Low Back Pain"[MeSH] OR "low back pain"[title]) AND ((((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[MeSH] AND "fusion"[title])) AND ("Treatment Outcome"[MeSH] OR "Patient Satisfaction"[MeSH] OR "functional outcome" [All Fields] OR "functional outcomes"[All Fields] OR "outcome"[title] OR "outcomes"[title] OR "clinical outcome"[All Fields] OR "clinical outcomes"[All Fields]) AND (("2003"[PDAT]: "3000"[PDAT]) AND "humans" [MeSH] AND English[lang])). Search results were limited to human studies, English language, and patients between the ages of 18 and 65. Duplicates, technical notes, reviews, and other publications that did not describe the use of lumbar fusion for patients with low-back pain without stenosis or spondylolisthesis were discarded. The bibliographies of the selected articles were inspected and additional relevant papers were identified. Three clinical series and one systematic review that contributed to the guideline formulation are described in Table 1. The remaining references provided additional background information and are included in the bibliography.

Scientific Foundation

A review of the Cochrane database failed to identify a randomized, controlled trial investigating the utility of lumbar fusion for the treatment of low-back pain due to spondylosis.[19] Two subsequent randomized trials were summarized in the original version of the Lumbar Fusion Guidelines.[29] Fritzell et al. performed a randomized, controlled, multicenter trial of patients presenting with back pain presumably due to lumbar spondylosis. A total of 294 patients were randomized to one of 3 surgical groups or to physical therapy.[16] At 2 years' follow-up, 289 (98%) of the initial 294 patients remained in the study, but 25 had changed treatment groups. Each of the surgical groups achieved better clinical outcomes than the conservatively treated cohort. Back pain was reduced by 33% in the surgical groups versus 7% in the control group ($p = 0.0002$). The Oswestry Disability Index (ODI) score improved by 25% in the surgically treated patients and only 6% in the controls ($p = 0.015$). The return-to-work rate was 36% in the surgically treated patients versus 13% in the controls ($p = 0.002$). The authors concluded that lumbar fusion was a more effective treatment option for patients with chronic low-back pain after failure of conservative care than traditional nonsurgical treatment. Limitations of this study included a lack of well-defined conservative treatment group and patient crossover. This study provides Level II evidence in favor of lumbar fusion over traditional nonoperative treatment for patients with low-back pain.

In a smaller study of 64 patients, Brox et al. compared instrumented fusion versus physical therapy with cognitive exercises in patients presenting with chronic low-back pain and spondylosis at L4–5 and/or L5–S1 on plain radiographs.[7] Patients were followed for 1 year with a 97% follow-up rate. The main outcomes measure was the ODI. The mean difference between the groups was 2.3 (not significant, $p = 0.33$). Limitations of the study included a small sample size and wide variation between patients. This study provides Level II data for the equivalence between lumbar fusion and physical and cognitive therapy for patients with low-back pain.

There have been 3 prospective, randomized trials comparing lumbar fusion to conservative treatment and one systematic review of the literature since the publication of the original guidelines.[6,9,14,26]

Brox et al. performed a prospective, randomized study comparing the clinical results of lumbar fusion versus cognitive intervention and exercises in patients with chronic back pain following surgery for lumbar disc herniation.[6] Inclusion criteria consisted of age 25–60 years, at least 1 year of low-back pain following a lumbar discectomy, and lumbar disc degeneration at L4–5 and/or L5–S1 on plain radiographs. Patients were excluded if there was evidence of spinal stenosis, widespread myofascial pain, recurrent disc herniations, inflammatory disease, fracture, previous lumbar fusion, and/or psychiatric disease. Patients were randomized to either posterolateral instrumented fusion with autologous bone graft or cognitive intervention and exercises. The rehabilitation program lasted approximately 25 hours per week for 3 weeks. Patients were given information on the relevant anatomy and the mechanisms of pain. They were instructed that they could not harm themselves during routine activities of daily living. The specific exercise program was tailored to the individual patient. The ODI was used as the primary outcome measure. Secondary measures included the VAS, medication usage, General Function Score, Global Back Disability Questionnaire, work status, and the Prolo scale. The General Function Score consists of 9 questions used to measure back-related disability in activities of daily living and has been previously validated.[23]

TABLE 1: Lumbar fusion versus conservative treatment for chronic low-back pain: summary of evidence*

Authors & Year	Level of Evidence	Description	Comments
Fritzell et al., 2001	II	Prospective, randomized, controlled, multicenter trial. 294 pts w/ LBP randomized to 1 of 3 surgical groups or PT (controls). At 2-yr follow-up 289 (98%) of the initial 294 pts remained, but 25 had changed treatment groups. Each surgical group had improved clinical outcomes compared to the PT cohort. Back pain was reduced by 33% in surgical groups vs 7% in controls (p = 0.0002). ODI improved 25% in the surgical groups vs 6% in controls (p = 0.015). Return-to-work rate was 36% in the surgical groups vs 13% in controls (p = 0.002). Authors concluded that lumbar fusion was more effective than traditional nonsurgical treatment for pts w/ chronic LBP after failure of conservative care.	Limitations of this study include a lack of a well-defined conservative treatment group & patient crossover. This study provides Level II evidence in favor of lumbar fusion over traditional nonoperative treatment for pts w/ LBP.
Brox et al., 2003	II	Prospective, randomized trial involving 64 pts & comparing instrumented fusion vs PT w/ cognitive exercises for chronic LBP. Pts were followed for 1 yr w/ a 97% follow-up rate. Outcomes measured w/ ODI. Mean difference btwn the groups was 2.3 (NS, p = 0.33).	Limitations of the study include a small sample size & wide variation btwn pts. This study provides Level II data for equivalence btwn lumbar fusion & PT w/ cognitive therapy for pts w/ LBP.
Fairbank et al., 2005	II	Prospective, randomized, multicenter trial. Outcome measures were the ODI & the shuttle walking test. Secondary outcomes included the SF-36, psychological assessment, complications, & work status. Pts evaluated at baseline & at 6, 12, & 24 mos. Lumbar fusion technique not defined & left to surgeon preference. Intensive rehabilitation program consisted of education & exercises 5 days a wk for 3 wks. Cognitive behavior therapy helped identify & overcome fears & unhelpful beliefs. 339 pts randomized across 15 centers. ODI improved from 46.5 to 34.0 in the surgical group vs 44.8 to 36.1 for control group. Improvement in surgical cohort was significant compared to the rehabilitation group (p = 0.045). Shuttle walking test also improved in both groups, but there were no significant differences. There were no statistically significant differences btwn the 2 groups in any of the secondary outcome measures.	Limitations of this study include a 20% loss to follow-up at 24 mos & 28% crossover from rehabilitation to surgery. The clinical significance of the difference in ODI scores is unclear, especially given the failure to observe a relevant difference in the other outcome measures. Based on these limitations, the study was downgraded to Level II evidence due to lack of benefit of lumbar fusion over intensive rehabilitation.
Brox et al., 2006	II	Prospective, randomized study comparing lumbar fusion vs cognitive intervention & exercises in pts w/ chronic back pain following surgery for lumbar disc herniation. Randomized to posterolateral instrumented fusion w/ autologous bone graft or cognitive intervention & exercises. Rehabilitation lasted approximately 25 hrs per wk for 3 wks. ODI was primary outcome measure. Secondary measures included VAS, medication usage, General Function Score, Global Back Disability Questionnaire, work status, & the Prolo Scale. Final study included 60 pts, 29 randomized to fusion & 31 to control. 97% follow-up rate at 1 yr, w/ 6 pts crossing over from surgery to the conservative treatment cohort & 2 pts in the conservative treatment group undergoing surgery. ODI significantly improved in both groups, from 47.0 to 38.1 for fusion & from 45.1 to 32.3 for control (p = 0.001). No significant difference btwn groups at final follow-up.	Limitations of this study include a small sample size despite the power analysis, a relatively brief follow-up period, & the study of a small subset of pts w/ chronic back pain who had a previous lumbar discectomy. The authors concluded that either treatment alternative may be considered for pts presenting w/ chronic LBP following discectomy. Based on these limitations, this study was downgraded to Level II evidence supporting the use of either lumbar fusion or intensive rehabilitation w/ cognitive therapy for the treatment of pts w/ chronic LBP w/out stenosis or spondylolisthesis.
Ohtori et al., 2011	II	Prospective study randomizing pts w/ discogenic LBP into surgical vs nonsurgical treatment groups. 41 pts w/ MRI evidence of disc degeneration at L4–5 or L5–S1 & pain provocation on discography w/ pain relief w/ discoblock. Surgery consisted of ALIF unless there was difficulty w/ anterior vessels, in which case a posterolateral instrumented fusion was performed. Nonsurgical control included daily walking & exercises following individual instruction. Outcomes measures included VAS, ODI, & JOA score. Subjective outcomes were graded according to the NASS Low Back Outcome Instrument. Radiographic evaluation of fusion was performed w/ AP radiographs & CT scan by 3 blinded observers. Data were compared preoperatively & at 1 & 2 yrs. All pts had solid arthrodesis. Each of the outcomes measures was significantly better for the surgical groups at 2 yrs as compared to the nonoperative group (p < 0.05).	Limitations of the study include a small sample size & inconsistency in the type of surgery performed. This study provides Level II evidence in favor of surgery over walking & exercises in pts w/ chronic discogenic LBP.

(continued)

TABLE 1: Lumbar fusion versus conservative treatment for chronic low-back pain: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Comments
Chou et al., 2009	II	Systematic review of the literature to assess risks & benefits of surgical treatment of back pain. 1449 citations were reviewed, which led to a review of 24 full-text articles on surgery for LBP for degenerative disorders, 4 of which compared surgery to nonsurgical therapy. Guidelines suggested fusion was no more effective than an intensive rehabilitation program, but fusion was associated w/ moderate benefits as compared to traditional conservative treatment options.	This systematic review provides Level II evidence for the equivalence of surgery to an intensive rehabilitation program & moderate benefits of surgery over traditional nonoperative treatment options.

* ALIF = anterior lumbar interbody fusion; AP = anteroposterior; JOA = Japanese Orthopaedic Association; LBP = low-back pain; NASS = North American Spine Society; NS = not significant; ODI = Oswestry Disability Index; PT = physical therapy; pts = patients; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale.

The authors initially performed a pilot study to determine the required sample size to demonstrate statistical significance. The final study consisted of 60 patients, 29 randomized to fusion and 31 receiving cognitive intervention and exercises. The percentage of male patients was significantly lower in the surgery group (38%) than in the cognitive treatment and exercise group (64%) ($p = 0.04$). There was a 97% follow-up rate at one-year, with 6 patients crossing over from surgery to the conservative cohort and 2 patients in the conservative treatment group undergoing surgery. The mean ODI scores improved significantly in both groups (from 47.0 to 38.1 in the fusion group and from 45.1 to 32.3 in the conservative treatment group, $p = 0.001$). There was no significant difference between the 2 groups at final follow-up. Limitations of this study included a small sample size despite the power analysis, a relatively brief follow-up period, and the study of a small subset of patients with chronic back pain who had a previous lumbar discectomy. The authors concluded that either treatment alternative may be considered for patients presenting with chronic low-back pain following discectomy. Based on these limitations this study was downgraded to Level II evidence in supporting the use of either lumbar fusion or intensive rehabilitation with cognitive therapy for the treatment of patients with chronic low-back pain without stenosis or spondylolisthesis.

A large, prospective, randomized, multicenter trial was performed by Fairbank et al. to assess the effectiveness of spinal fusion versus an intensive rehabilitation program for patients with chronic low-back pain.[14] All patients were between 18 and 55 years of age and had at least a 1-year history of low-back pain. Exclusion criteria consisted of infection, inflammatory disease, tumor, fracture, psychiatric disorders, pregnancy, and previous spinal fusion. Outcome measures were the ODI and the shuttle walking test.[31] Secondary outcomes included the 36-Item Short Form Health Survey (SF-36), psychological assessment, complications, and work status. Patients were evaluated at baseline and after 6, 12, and 24 months. The technique of the lumbar fusion was not defined and was left to surgeon preference. The intensive rehabilitation program consisted of education and exercises 5 days each week for 3 consecutive weeks. The program was individually tailored to each patient and modified based on the patient's response. Cognitive behavior therapy was also included to help identify and overcome fears and unhelpful beliefs. This program educated patients on their anatomy and causes of pain and encouraged them to perform normal activities of daily living that would not cause them harm.[17]

The study population consisted of 339 patients randomized across 15 different centers. There were no significant baseline differences between the 2 cohorts. A significant crossover rate was observed in the conservatively treated patients. Forty-eight (28%) of the patients randomized to the intensive rehabilitation group eventually had surgery; however, only 7 (4%) of the patients randomized to surgery were treated with rehabilitation alone. There were 19 surgical complications, of which 11 required additional surgery. There were no complications attributed to the rehabilitation program. The mean ODI scores improved from 46.5 to 34.0 in the surgical group versus 44.8 to 36.1 for the rehabilitation group. The extent of improvement observed in the surgical cohort proved to be significant when compared with the outcome within the rehabilitation group ($p = 0.045$). Performance on the shuttle walking test also improved in both groups, but there were no significant differences between the 2 cohorts. There were no statistically significant differences between the 2 groups in any of the secondary outcome measures. Limitations of this study included a 20% loss to follow-up at 24 months and 28% crossover from rehabilitation to surgery. The clinical significance of the difference in ODI scores is unclear, especially given the failure to observe a relevant difference in the other outcome measures. Based on these limitations, the study was downgraded to Level II evidence because of lack of benefit of lumbar fusion over intensive rehabilitation.

Ohtori et al. performed a prospective study randomizing patients with discogenic low-back pain into surgical versus nonsurgical treatment groups.[26] The study consisted of 41 patients with MRI evidence of disc degeneration at either L4–5 or L5–S1 and pain provocation on discography with pain relief with discoblock (a procedure that involves injecting the disc with an anesthetic agent instead of a contrast agent in an effort to eliminate as opposed to reproducing a patient's pain). Surgery consisted of anterior lumbar interbody fusion unless there was presumed to be difficulty with the anterior vessels, in which case a posterolateral instrumented fusion was performed. The nonsurgical control group included daily walking and exercises following

individual instruction. Outcomes measures included VAS, ODI, and the JOA score. Subjective outcomes were graded according to the North American Spine Society Low Back Outcome Instrument. Radiographic evaluation of fusion was performed with anteroposterior radiographs and CT scan by 3 blinded observers. Data were compared preoperatively and at 1 and 2 years postoperatively.

All of the patients undergoing surgery went on to achieve a solid arthrodesis. Each of the outcomes measures were significantly better for the 2 surgical groups at 2 years as compared with the nonoperative group ($p < 0.05$). Limitations of the study included a small sample size and inconsistency in the type of surgery performed. This study provides Level II evidence in favor of surgery over walking and exercises in patients with chronic discogenic low-back pain.

Several other recent studies have provided additional data on outcomes of lumbar fusion in patients with chronic low-back pain.[2–5,8,10–13,15,18,20–22,24,25,27,28,30,33,34] These studies did not have a randomized control group receiving nonoperative treatment. Instead they compared different fusion techniques, compared fusion versus arthroplasty, or failed to include an adequate control group. As a result, they were classified as case series and only provided Level IV evidence supporting the utility of lumbar fusion for the treatment of patients with chronic back pain.

Summary

The results of this review reveal a lack of sufficient evidence to support a single treatment alternative in patients with intractable low-back pain without stenosis or spondylolisthesis. There is Level II evidence supporting the use of either intensive rehabilitation programs with a cognitive component or lumbar fusion. No significant clinical difference in outcomes were observed between these 2 options, but such rehabilitation programs are not generally available in most areas. There is Level II evidence to support lumbar fusion over traditional physical therapy alone, but that benefit is not present when fusion is compared with a more intensive physical therapy program with cognitive therapy. Numerous case series, constituting Level IV evidence, support the use of lumbar fusion in this patient population. These studies reported significant clinical improvements but failed to incorporate an adequate control group for comparison purposes.

Key Issues for Future Investigation

Treatment of intractable low-back pain in patients without stenosis or spondylolisthesis remains a difficult problem with many unanswered questions. Further investigation will be necessary to improve the diagnostic capabilities of identifying the origin of pain in this patient population. With improved diagnostic capabilities, intervention can be directed at the primary pathological process.

While there is currently a lack of high-quality Level I evidence to support the use of lumbar fusion for these patients, there are numerous studies that demonstrate a definitive clinical improvement following fusion. Future investigation will be necessary in an attempt to identify factors, both patient specific and surgery related, that are predictive of outcome so that a subset of patients are defined who will respond favorably to fusion as compared with conservative management.

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., 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 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: Eck. 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: Eck. Study supervision: Kaiser.

References

1. Abdu WA, Lurie JD, Spratt KF, Tosteson AN, Zhao W, Tosteson TD, et al: Degenerative spondylolisthesis: does fusion method influence outcome? Four-year results of the spine patient outcomes research trial. **Spine (Phila Pa 1976) 34:**2351–2360, 2009
2. Anderson PA, Schwaegler PE, Cizek D, Leverson G: Work status as a predictor of surgical outcome of discogenic low back pain. **Spine (Phila Pa 1976) 31:**2510–2515, 2006
3. Arnold PM, Robbins S, Paullus W, Faust S, Holt R, McGuire R: Clinical outcomes of lumbar degenerative disc disease treated with posterior lumbar interbody fusion allograft spacer: a prospective, multicenter trial with 2-year follow-up. **Am J Orthop 38:**E115–E122, 2009
4. Aryan HE, Lu DC, Acosta FL Jr, Ames CP: Stand-alone anterior lumbar discectomy and fusion with plate: initial experience. **Surg Neurol 68:**7–13, 2007
5. Berg S, Tullberg T, Branth B, Olerud C, Tropp H: Total disc replacement compared to lumbar fusion: a randomised controlled trial with 2-year follow-up. **Eur Spine J 18:**1512–1519, 2009
6. Brox JI, Reikerâs O, Nygaard Ø, Sørensen R, Indahl A, Holm

I, et al: Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. **Pain 122:**145–155, 2006
7. Brox JI, Sørensen R, Friis A, Nygaard Ø, Indahl A, Keller A, et al: Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. **Spine (Phila Pa 1976) 28:**1913–1921, 2003
8. Burkus JK, Gornet MF, Schuler TC, Kleeman TJ, Zdeblick TA: Six-year outcomes of anterior lumbar interbody arthrodesis with use of interbody fusion cages and recombinant human bone morphogenetic protein-2. **J Bone Joint Surg Am 91:**1181–1189, 2009
9. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD: Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. **Spine (Phila Pa 1976) 34:**1094–1109, 2009
10. Dawson E, Bae HW, Burkus JK, Stambough JL, Glassman SD: Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. **J Bone Joint Surg Am 91:**1604–1613, 2009
11. Delamarter R, Zigler JE, Balderston RA, Cammisa FP, Goldstein JA, Spivak JM: Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement compared with circumferential arthrodesis for the treatment of two-level lumbar degenerative disc disease: results at twenty-four months. **J Bone Joint Surg Am 93:**705–715, 2011
12. Deutsch H, Musacchio MJ Jr: Minimally invasive transforaminal lumbar interbody fusion with unilateral pedicle screw fixation. **Neurosurg Focus 20(3):**E10, 2006
13. Esposito P, Pinheiro-Franco JL, Froelich S, Maitrot D: Predictive value of MRI vertebral end-plate signal changes (Modic) on outcome of surgically treated degenerative disc disease. Results of a cohort study including 60 patients. **Neurochirurgie 52:**315–322, 2006
14. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R, et al: Randomised controlled trial to compare surgical stabilization of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. **BMJ 330:**1233, 2005 (Erratum in **BMJ 330:**1485, 2005)
15. Fogel GR, Toohey JS, Neidre A, Brantigan JW: Is one cage enough in posterior lumbar interbody fusion: a comparison of unilateral single cage interbody fusion to bilateral cages. **J Spinal Disord Tech 20:**60–65, 2007
16. Fritzell P, Hägg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study Group: 2001 Volvo Award Winner in Clinical Studies. Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. **Spine (Phila Pa 1976) 26:**2521–2532, 2001
17. Frost H, Lamb SE, Shackleton CH: A functional restoration programme for chronic low back pain: a prospective outcome study. **Physiotherapy 86:**285–293, 2000
18. Gepstein R, Shabat S, Reichel M, Pikarsky I, Folman Y: Treatment of postdiscectomy low back pain by percutaneous posterior lumbar interbody fusion versus open posterior lumbar fusion with pedicle screws. **Spine J 8:**741–746, 2008
19. Gibson JN, Waddell G: Surgery for degenerative lumbar spondylosis: updated Cochrane Review. **Spine (Phila Pa 1976) 30:**2312–2320, 2005
20. Glassman SD, Carreon LY, Djurasovic M, Dimar JR, Johnson JR, Puno RM, et al: Lumbar fusion outcomes stratified by specific diagnostic indication. **Spine J 9:**13–21, 2009
21. Gornet MF, Burkus JK, Dryer RF, Peloza JH: Lumbar disc arthroplasty with Maverick disc versus stand-alone interbody fusion: a prospective, randomized, controlled, multicenter investigational device exemption trial. **Spine (Phila Pa 1976) 36:**E1600–E1611, 2011
22. Guyer RD, McAfee PC, Banco RJ, Bitan FD, Cappuccino A, Geisler FH, et al: Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: five-year follow-up. **Spine J 9:**374–386, 2009
23. Hägg O, Fritzell P, Odén A, Nordwall A, Swedish Lumbar Spine Study Group: Simplifying outcome measurement: evaluation of instruments for measuring outcome after fusion surgery for chronic low back pain. **Spine (Phila Pa 1976) 27:**1213–1222, 2002
24. Niemeyer T, Halm H, Hackenberg L, Liljenqvist U, Bövingloh AS: Post-discectomy syndrome treated with lumbar interbody fusion. **Int Orthop 30:**163–166, 2006
25. Ohtori S, Koshi T, Yamashita M, Takaso M, Yamauchi K, Inoue G, et al: Single-level instrumented posterolateral fusion versus non-instrumented anterior interbody fusion for lumbar spondylolisthesis: a prospective study with a 2-year follow-up. **J Orthop Sci 16:**352–358, 2011
26. Ohtori S, Koshi T, Yamashita M, Yamauchi K, Inoue G, Suzuki M, et al: Surgical versus nonsurgical treatment of selected patients with discogenic low back pain: a small-sized randomized trial. **Spine (Phila Pa 1976) 36:**347–354, 2011
27. Peng B, Chen J, Kuang Z, Li D, Pang X, Zhang X: Diagnosis and surgical treatment of back pain originating from endplate. **Eur Spine J 18:**1035–1040, 2009
28. Rapan S, Jovanović S, Gulan G: Transforaminal lumbar interbody fusion (TLIF) and unilateral transpedicular fixation. **Coll Antropol 34:**531–534, 2010
29. 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 7: intractable low-back pain without stenosis or spondylolisthesis. **J Neurosurg Spine 2:**670–672, 2005
30. Suratwala SJ, Pinto MR, Gilbert TJ, Winter RB, Wroblewski JM: Functional and radiological outcomes of 360 degrees fusion of three or more motion levels in the lumbar spine for degenerative disc disease. **Spine (Phila Pa 1976) 34:**E351–E358, 2009
31. Taylor S, Frost H, Taylor A, Barker K: Reliability and responsiveness of the shuttle walking test in patients with chronic low back pain. **Physiother Res Int 6:**170–178, 2001
32. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al: Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. **N Engl J Med 356:**2257–2270, 2007
33. Zhou ZJ, Zhao FD, Fang XQ, Zhao X, Fan SW: Meta-analysis of instrumented posterior interbody fusion versus instrumented posterolateral fusion in the lumbar spine. A review. **J Neurosurg Spine 15:**295–310, 2011
34. Zigler J, Delamarter R, Spivak JM, Linovitz RJ, Danielson GO III, Haider TT, et al: Results of the prospective, randomized, multicenter, Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. **Spine (Phila Pa 1976) 32:**1155–1163, 2007

Manuscript submitted March 13, 2014.
Accepted April 4, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14270.
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.

J Neurosurg Spine 21:48–53, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 8: Lumbar fusion for disc herniation and radiculopathy

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

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

Patients suffering from a lumbar herniated disc will typically present with signs and symptoms consistent with radiculopathy. They may also have low-back pain, however, and the source of this pain is less certain, as it may be from the degenerative process that led to the herniation. The surgical alternative of choice remains a lumbar discectomy, but fusions have been performed for both primary and recurrent disc herniations. In the original guidelines, the inclusion of a fusion for routine discectomies was not recommended. This recommendation continues to be supported by more recent evidence. Based on low-level evidence, the incorporation of a lumbar fusion may be considered an option when a herniation is associated with evidence of spinal instability, chronic low-back pain, and/or severe degenerative changes, or if the patient participates in heavy manual labor. For recurrent disc herniations, there is low-level evidence to support the inclusion of lumbar fusion for patients with evidence of instability or chronic low-back pain. *(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14271)*

Key Words • **fusion** • **lumbar spine** • **herniated disc** • **practice guidelines**

Recommendations

There is no evidence that conflicts with the previous recommendations formulated from the first generation of the Lumbar Fusion Guidelines.

Grade C

Lumbar spinal fusion is not recommended as a routine treatment following primary disc excision in patients with isolated herniated lumbar discs causing radiculopathy (Level IV evidence).

Lumbar spinal fusion is a potential option in patients with herniated discs who have evidence of significant chronic axial back pain, work as manual laborers, have severe degenerative changes, or have instability associated with radiculopathy caused by herniated lumbar discs (Level IV evidence).

Reoperative discectomy and fusion is a treatment option in patients with recurrent disc herniations associated with instability or chronic axial low back pain (Level III and IV evidence).

Rationale

Herniation of a lumbar disc will typically manifest with radicular signs and symptoms consistent with the spinal nerve under compression. Less specific complaints of low-back pain, presumably from the degeneration associated with the disc herniation, may also be present. To address the primary pathology, which is the compression

Abbreviations used in this paper: JOA = Japanese Orthopaedic Association; LHNP = lumbar herniated nucleus pulposus; ODI = Oswestry Disability Index.

of the spinal nerve, lumbar discectomy has become the established surgical procedure in cases in which conservative management fails to provide relief.

Although spinal fusion is routinely performed for lumbar instability or low-back pain associated with severe disc degeneration, it has been used for patients presenting with either a primary or recurrent lumbar herniated nucleus pulposus (LHNP) and this application has been described in the literature. Incorporating a fusion during a routine discectomy would increase the complexity of the case, prolong the surgical time, and potentially increase complication rates, without proven medical necessity. As indicated in the first generation of the Lumbar Fusion Guidelines, justification for fusion under these circumstances is lacking.[15] The purpose of this update is to examine the more recent literature investigating the role of fusion in the operative management of patients presenting with radiculopathy and/or back pain secondary to a LHNP.

Search Criteria

A computerized search of the database of the National Library of Medicine from July 2003 to December 2011 was conducted using the search terms ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ("Radiculopathy"[MeSH] OR radiculopathy[title] OR "intervertebral disk displacement"[title] OR "herniated"[title] OR "intervertebral disc displacement"[title] OR "herniation"[title]) AND (("2003"[PDAT]: "3000"[PDAT]) AND "humans"[MeSH] AND English[lang]). The search was restricted to the English language. This yielded a total of 74 references. The titles and abstracts of each of these references were reviewed, and papers not concerned with the use of fusion with lumbar disc herniations were discarded. References were identified that provided either direct or supporting evidence relevant to the use of fusion as a treatment for lumbar disc herniations. These papers were obtained and reviewed, and relevant references from the bibliographies of these papers were identified. Relevant papers providing Level IV or better evidence are summarized in the evidentiary table. Other papers providing supportive data are shown in the reference section.

Scientific Foundation

Primary Herniated Disc With Radiculopathy

In the first generation of the Lumbar Fusion Guidelines, Resnick et al. examined the role of fusion for patients with radiculopathy and an LHNP.[15] The authors performed a literature review of studies of Level IV or better quality and determined that the routine use of fusion in conjunction with a disc excision for primary LHNP is not recommended. The outcome following decompressive surgery for a patient presenting with an LHNP and radiculopathy, whether primary or recurrent, has been demonstrated in numerous publications.[10,11] There are a plethora of studies reporting excellent results and outcomes for patients with primary disc herniations having decompressive surgeries without fusion, and many of these studies are Level I and II studies.[1–4,14,17,20–23]

This current review will examine the studies investigating fusion as compared with discectomy alone to determine if evidence exists for the addition of fusion in patients with primary disc herniation. Advocates for fusion during the index discectomy claim that stabilizing the segment may prevent late-onset instability and the development of chronic low-back pain. Although several studies have demonstrated that the occurrence of instability following discectomy is associated with less-favorable outcomes, the incidence is relatively low, and therefore routine fusion is not recommended.[15,16]

Some of the studies used to support this recommendation in the past review were examined once again. Takeshima et al. performed a retrospective review of cases involving patients undergoing surgery for primary disc herniations.[19] Of 95 patients, 44 underwent discectomy alone (after 1990) and 51 underwent discectomy and fusion (between 1986 and 1989), with follow-up averaging 7 years and assessments using the Japanese Orthopaedic Association (JOA) rating scale. There was no statistically significant difference between the 2 groups ($p = 0.31$). This study provides Level III evidence that the routine use of a noninstrumented posterolateral fusion does not improve functional outcome in patients treated with lumbar discectomy. In another study, Donceel and DuBois reviewed a series of 3956 cases involving patients with disc herniations treated with either discectomy (n = 3670) or discectomy and fusion (n = 286).[6] They found that 70% of the discectomy-alone group were able to resume their preoperative work level at 1 year after surgery, compared with 45% of the fusion group. The authors noted that the fusion group tended to have more significant symptoms and more complex preoperative histories. This retrospective review provides Level IV evidence suggesting that the addition of fusion does not improve patient outcomes. There were no further studies found that compared discectomy alone to discectomy and fusion. This is likely due to the large number of studies demonstrating excellent outcomes without fusion in this patient population with an isolated LHNP.

Primary Herniated Lumbar Disc Associated With Low-Back Pain/Instability

Fusion has also been recommended for patients presenting with new-onset LHNP and radiculopathy in the presence of axial low-back pain or radiographic instability. The previous review concluded that there was Level III evidence to support the use of posterior fusion at the time of initial discectomy surgery in manual laborers or those with significant preoperative axial low-back pain.[15] Matsunaga et al. reported the results of a retrospective study of 80 cases involving manual laborers treated via either open or percutaneous discectomy (n = 51) or an open discectomy and fusion (n = 29).[12] They found that at the 1-year point 53% of the patients in the discectomy group and 89% of those in the fusion group were able to resume and maintain preoperative manual labor work activities. Although the discectomy patients did return to

work earlier (12 weeks after surgery) than those in the fusion group (25 weeks after surgery), 22% of the discectomy group could not maintain their work activities due to "lumbar fatigue." These authors concluded that the addition of fusion should be considered in manual laborers, as it seems to provide a better chance of returning to and staying at their preoperative level of function. The paper is judged to provide Level IV evidence supporting the use of posterolateral fusion at the time of discectomy to improve return to work rates in patients involved in heavy manual labor work activities (see Table 1).

Eie reported on 259 patients with disc herniations who were treated either by discectomy alone (n = 119) or by discectomy and noninstrumented posterolateral fusion (n = 68).[7] At 6 years postsurgery, 76% of the discectomy-alone group reported satisfaction compared with 85% of the fusion group. The discectomy-alone patients reported a significantly higher incidence of pain recurrence (27% of patients) compared with the discectomy plus fusion group (15% of patients, $p < 0.01$). This is another Level IV paper supporting the use of fusion at the time of discectomy, especially in patients with significant low-back pain, as they have a higher chance of having pain in later years without a concomitant fusion (see Table 1).

Newer studies reviewed since the prior recommendations were published support the use of fusion for patients with significant preoperative low-back pain and those with existing instability. Satoh et al. published a retrospective review of 174 cases involving patients with disc herniations treated with fusion and 177 involving patients treated with discectomy alone.[18] All patients had at least 5 years of follow-up. Fusion criteria included either a massive disc herniation, as defined by a complete myelographic block on a CT myelogram, or segmental instability, as defined by an anterolisthesis of greater than 3 mm with or without local kyphosis of more than 5° on a flexion lateral radiograph. Patients were assessed on a clinical outcomes questionnaire with a scale consisting of excellent, good, fair, and poor, which appeared to be a modification of Odom's criteria. Patients undergoing a fusion demonstrated significantly better outcomes with respect to low-back pain. The frequency of revision surgery was significantly higher in patients who did not receive a fusion, but met the criteria for fusion. Interestingly, those patients who did not fulfill the criteria for fusion but had a fusion surgery also had significantly better results in terms of low-back pain scores compared with those without fusions. The authors concluded with this Level IV study that patients with disc herniations and instability or massive herniations can be successfully treated with fusion at the time of primary discectomy.

Recurrent Disc Herniation

The previous Lumbar Fusion Guidelines concluded that reoperative discectomy is recommended as a treatment option in patients with recurrent disc herniations and radiculopathy.[15] For a first-time recurrence, this recommendation continues to be supported by more recent publications. Fu et al., in a retrospective Level III review, investigated the outcome in 41 cases of recurrent lumbar disc herniation.[8] In this study, 23 patients underwent a revision discectomy and 18 underwent a revision discectomy with posterolateral instrumentation and fusion.[8] The minimum follow-up for both cohorts was 60 months, and patients were evaluated using the JOA scores for low-back pain. The clinical outcome was excellent or good in 78.3% of the discectomy cohort and 83.3% of the fusion group. There was no significant difference in clinical outcome parameters between the 2 groups including low-back pain scores, but intraoperative blood loss, length of surgery, and length of hospitalization were significantly less in the nonfusion group. This study provides Level III evidence that in patients presenting with an isolated recurrent herniation with sciatica, disc excision alone without fusion is recommended. This study had very few patients lost to follow-up and was from a single-center with excellent longer-term follow-up.

Fusion at the time of revision discectomy has been more consistently recommended as a treatment option for patients with associated lumbar instability, radiographic degenerative changes, and/or chronic axial low-back pain.[15] This recommendation in the previous guidelines was based on several studies. Huang and Chen reported on 28 patients undergoing posterior interbody fusion (8 with recurrent disc herniations and 10 with low-grade spondylolisthesis).[9] These patients all had significant degenerative changes and some had spondylolisthesis. The average follow-up was 14 months, and all patients had pedicle screw fixation. Overall, 93% of the patients were satisfied with their condition, and 82% were considered to have achieved radiographic fusion. Of the 8 patients with recurrent disc herniations, 6 had excellent or good outcomes, and 2 had fair outcomes. In another study, Chitnavis and colleagues reported on a group of patients with recurrent disc herniations with symptoms of back pain or signs of instability, who were treated with posterior decompression and interbody fusion.[5] Of a total of 50 patients with 6 months to 5 years of follow-up, 92% improved after surgery and 90% were very satisfied with their results. This study provides Level IV evidence demonstrating good results with fusion in these patients with recurrent disc herniations with instability and/or axial low-back pain. There is a paucity of more recent evidence to support or refute the previous conclusions from the initial publication of the Lumbar Fusion Guidelines. The majority of these studies were case series and not comparative studies looking at discectomy alone versus discectomy and fusion.[13]

Summary

Based on the recent literature reviewed, there does not appear to be evidence to support the routine use of fusion at the time of an index discectomy operation. There remains conflicting Level III and IV evidence regarding the potential benefit of the addition of fusion in certain situations; however, the increase in morbidity, cost, and potential complications associated with the use of fusion are not justified in routine situations. Patients with demonstrated preoperative instability and significant chronic low-back pain in addition to radicular symptoms may be

TABLE 1: Lumbar fusion for disc herniation and radiculopathy: summary of evidence*

Authors & Year	Level of Evidence	Brief Description	Comment
Eie, 1978	IV	A retrospective study of 259 pts: 119 discectomy only & 68 discectomy w/ in situ noninstrumented PLF. At the 6-mo follow-up, 89% of the discectomy-only pts & 88% of the fusion pts were satisfied. At the 6-yr follow-up, satisfaction rates were 76% vs 85% (NS). The rate of pain recurrence was much higher in the discectomy group (27% vs 15%, $p < 0.01$). Ability to maintain work at the preop status was 79% in the discectomy group & 86% in the fusion group at 6 yrs.	Results were similar at 6 mos, but fusion provided more stable relief of pain & maintenance of work function over prolonged follow-up.
Matsunaga et al., 1993	IV	A retrospective review of 82 pts (defined as laborers) who engaged daily in work that repeatedly put large amounts of stress on their lumbar spine & 28 athletes. 81 pts had discectomy alone (microdiscectomy in 30 & percutaneous discectomy in 51); 29 pts had discectomy & fusion. The groups had similar demographics w/ a slightly shorter duration of Sx in the percutaneous discectomy group. Follow-up range: 2–7 yrs. Return to work at 1 yr: 75% in discectomy group, but 22% could not sustain work, leaving 53% in end; 89% in spinal fusion group; 58% in percutaneous discectomy group. Time to return to work: 9 wks for percutaneous discectomy, 15 wks for microdiscectomy, 25 wks for fusion. "Lumbar fatigue" given as reason pts w/ discectomy alone had to stop their work or change their jobs from manual labor.	Manual laborers do better after spinal fusion. The selection criteria for defining which pts are appropriate for fusion is unclear & validated outcome assessment was not used.
Donceel & Du Bois, 1998	IV	A retrospective series of 3956 cases (percutaneous nucleotomy in 126, microdiscectomy in 3544, discectomy & fusion in 286) taken from a Belgian sickness fund database. Follow-up averaged 1–3 yrs & fitness to resume work w/in 12 mos was seen in 70% of the discectomy group but only 45% of the fusion group ($p < 0.0001$). The poorest overall outcomes were seen in the fusion group: unemployment, a longer work capacity prior to surgery, longer hospital stay, older age, blue collar job, & lower compensation. The authors found no differences btwn the 2 discectomy techniques.	Large cohort w/ clearly poorer outcomes w/ fusion, but no validated outcome measures were used. In addition, there were no clear surgical indications described.
Takeshima et al., 2000	III	A retrospective study evaluating clinical & radiographic results in 95 pts w/ lumbar disc herniation. The purpose was to evaluate the results of disc excision, w/ & w/o posterolateral fusion. 44 pts underwent disc excision (after 1990), & 51 pts underwent disc excision & fusion (between 1986 & 1989). Clinical Sx were evaluated using the JOA back scores. Clinical outcome was excellent or good in 73% of the nonfusion group & in 82% of the fusion group ($p = 0.31$). The fusion group had greater postop reduction in LBP & lower rate of recurrence.	Retrospective study looking at discectomy vs discectomy & fusion for primary disc herniation. The authors concluded that there is seldom an indication for primary fusion for an isolated lumbar disc herniation.
Chitnavis et al., 2001	IV	A retrospective study of 50 cases of recurrent lumbar disc herniation w/ both leg pain & LBP treated w/ posterior interbody spinal fusion & cage placement; follow-up 6 mos–5 yrs. Pts in whom MRI demonstrated "simple" recurrent herniation w/ no back pain & those w/ only LBP were all excluded from the study. Surgery was performed in pts w/ Sx of neural root compression, tension signs, & back pain w/ focal disc degeneration & nerve root distortion on MRI compatible w/ clinical signs & Sx. In 40 pts (80%) pedicle screws were not used. Clinical outcome was assessed using the Prolo scale. 46 pts (92%) had Sx improvement at 6-mo follow-up, & at the latest follow-up, 45 (90%) would have undergone the same surgery again. Two-thirds of pts experienced good or excellent outcomes (Prolo score ≥ 8) at early & late follow-up. The fusion rate at 2 yrs after surgery was 95%. Only 36 of the 50 pts had more than 2 yrs of follow-up.	Pts w/ recurrent herniated discs w/ degenerative changes & LBP were treated successfully w/ posterior fusion. The authors felt that fusion was a good treatment option for pts w/ recurrent disc herniation w/ both leg pain & LBP w/ focal degenerative changes.
Fu et al., 2005	III	Retrospective study of 41 pts w/ repeat surgery for recurrent lumbar disc herniation w/ & w/o lumbar fusion. The mean follow-up was 88.7 mos (range 60–134). Clinical outcomes were excellent or good in 80.5% of pts (78.3% w/ discectomy alone & 83.3% w/ fusion; NS). No differences in LBP scores (2.2 in nonfusion group & 2.2 in fusion group at final follow-up) & no preop differences in scores or demographics/Sx.	The authors concluded that the results of fusion vs repeat discectomy alone were not significantly different, although there were slightly better results in the fusion group. Discectomy w/o fusion is recommended.

(continued)

TABLE 1: Lumbar fusion for disc herniation and radiculopathy: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Brief Description	Comment
Satoh et al., 2006	IV	Retrospective review of 174 pts w/ disc herniation treated w/ fusion & 177 treated w/ discectomy alone. All pts had at least 5 yrs of follow-up. Criteria for fusion were either massive disc herniation (complete myelographic block on CT/myelogram) or segmental instability (defined as an anterior slip of 3 mm &/or local kyphosis of >5° on a lateral flexion radiograph). These pts were divided into 4 groups (Group 1, fusion indicated & performed; Group 2, fusion not indicated but performed; Group 3, fusion indicated but not performed; & Group 4, fusion not indicated & not performed). The pts who had lumbar fusion had statistically superior results compared to the pts undergoing discectomy alone in terms of Sx of LBP, regardless of whether there was an indication for fusion ($p < 0.05$). The worst outcome in terms of LBP & leg pain scores was in the group in which fusion was indicated but only discectomy was performed (w/o fusion). In this group, the frequency of additional surgery was significantly higher. No validated outcomes scores were used to evaluate these cases.	Pts treated w/ fusion compared to those w/ discectomy alone have higher success rates in terms of Sx of LBP. Lumbar disc herniation w/ massive herniations or segmental instability can be well treated w/ posterior interbody fusion.
Huang & Chen, 2003	IV	28 pts (8 w/ recurrent disc herniations & 20 w/ low-grade degenerative spondylolisthesis) were treated w/ posterior decompression & interbody fusion w/ placement of a single threaded titanium cage w/ pedicle screw supplementation. The mean follow-up was 14.4 mos (range 8–39). Clinical outcomes were assessed using the Prolo scale. Dynamic radiography for fusion mass was interpreted by an independent radiologist. Overall, 92.86% of the pts were satisfied w/ their condition after surgery. Radiography showed the rate of bony fusion as 82.14%. All pts had significant degenerative disc disease w/ some having spondylolisthesis.	The authors concluded that recurrent disc herniations w/ LBP & degenerative changes can be managed w/ spinal fusion.

* LBP = low-back pain; NS = not significant; PLF = posterolateral lumbar fusion; Prolo scale = Prolo Functional Economic Outcome Rating Scale; pts = patients; Sx = symptoms.

candidates for fusion at the time of primary disc excision. Patients with recurrent disc herniations have been treated successfully with repeated excision as well as with excision and fusion. In patients with significant spinal deformity, instability, or associated chronic low-back pain, consideration of fusion is reasonable.

Key Issues for Future Investigation

The fact that fusion surgery is not required following a routine, index discectomy is well established, but further investigation to define various radiographic findings predicative of progressive disease would be very valuable. The utility of fusion for recurrent disc herniation remains controversial, and further investigation incorporating improved study design will be required to address this issue.

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., 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 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: Wang. 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: Wang. Study supervision: Kaiser.

References

1. Anderson PA, McCormick PC, Angevine PD: Randomized controlled trials of the treatment of lumbar disk herniation: 1983-2007. **J Am Acad Orthop Surg 16:**566–573, 2008

2. Arts MP, Brand R, van den Akker ME, Koes BW, Bartels RH, Tan WF, et al: Tubular diskectomy vs conventional microdiskectomy for the treatment of lumbar disk herniation: 2-year results of a double-blind randomized controlled trial. **Neurosurgery 69:**135–144, 2011
3. Atlas SJ, Keller RB, Wu YA, Deyo RA, Singer DE: Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the Maine Lumbar Spine Study. **Spine (Phila Pa 1976) 30:**927–935, 2005
4. Bruggeman AJ, Decker RC: Surgical treatment and outcomes of lumbar radiculopathy. **Phys Med Rehabil Clin N Am 22:**161–177, 2011
5. Chitnavis B, Barbagallo G, Selway R, Dardis R, Hussain A, Gullan R: Posterior lumbar interbody fusion for revision disc surgery: review of 50 cases in which carbon fiber cages were implanted. **J Neurosurg 95 (2 Suppl):**190–195, 2001
6. Donceel P, Du Bois M: Fitness for work after surgery for lumbar disc herniation: a retrospective study. **Eur Spine J 7:**29–35, 1998
7. Eie N: Comparison of the results in patients operated upon for ruptured lumbar discs with and without spinal fusion. **Acta Neurochir (Wien) 41:**107–113, 1978
8. Fu TS, Lai PL, Tsai TT, Niu CC, Chen LH, Chen WJ: Long-term results of disc excision for recurrent lumbar disc herniation with or without posterolateral fusion. **Spine (Phila Pa 1976) 30:**2830–2834, 2005
9. Huang KF, Chen TY: Clinical results of a single central interbody fusion cage and transpedicle screws fixation for recurrent herniated lumbar disc and low-grade spondylolisthesis. **Chang Gung Med J 26:**170–177, 2003
10. Kast E, Oberle J, Richter HP, Börm W: Success of simple sequestrectomy in lumbar spine surgery depends on the competence of the fibrous ring: a prospective controlled study of 168 patients. **Spine (Phila Pa 1976) 33:**1567–1571, 2008
11. Kim JS, Lee SH, Moon KH, Lee HY: Surgical results of the oblique paraspinal approach in upper lumbar disc herniation and thoracolumbar junction. **Neurosurgery 65:**95–99, 2009
12. Matsunaga S, Sakou T, Taketomi E, Ijiri K: Comparison of operative results of lumbar disc herniation in manual laborers and athletes. **Spine (Phila Pa 1976) 18:**2222–2226, 1993
13. Niemeyer T, Halm H, Hackenberg L, Liljenqvist U, Bövingloh AS: Post-discectomy syndrome treated with lumbar interbody fusion. **Int Orthop 30:**163–166, 2006
14. Pearson AM, Blood EA, Frymoyer JW, Herkowitz H, Abdu WA, Woodward R, et al: SPORT lumbar intervertebral disk herniation and back pain: does treatment, location, or morphology matter? **Spine (Phila Pa 1976) 33:**428–435, 2008
15. 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 8: lumbar fusion for disc herniation and radiculopathy. **J Neurosurg Spine 2:**673–678, 2005
16. Resnick DK, Groff MC: Evidence-based guidelines in lumbar spine surgery. **Prog Neurol Surg 19:**123–134, 2006
17. Rihn JA, Hilibrand AS, Radcliff K, Kurd M, Lurie J, Blood E, et al: Duration of symptoms resulting from lumbar disc herniation: effect on treatment outcomes: analysis of the Spine Patient Outcomes Research Trial (SPORT). **J Bone Joint Surg Am 93:**1906–1914, 2011
18. Satoh I, Yonenobu K, Hosono N, Ohwada T, Fuji T, Yoshikawa H: Indication of posterior lumbar interbody fusion for lumbar disc herniation. **J Spinal Disord Tech 19:**104–108, 2006
19. Takeshima T, Kambara K, Miyata S, Ueda Y, Tamai S: Clinical and radiographic evaluation of disc excision for lumbar disc herniation with and without posterolateral fusion. **Spine (Phila Pa 1976) 25:**450–456, 2000
20. Tosteson AN, Skinner JS, Tosteson TD, Lurie JD, Andersson GB, Berven S, et al: The cost effectiveness of surgical versus nonoperative treatment for lumbar disc herniation over two years: evidence from the Spine Patient Outcomes Research Trial (SPORT). **Spine (Phila Pa 1976) 33:**2108–2115, 2008
21. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, et al: Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. **JAMA 296:**2451–2459, 2006
22. Weinstein JN, Lurie JD, Tosteson TD, Tosteson AN, Blood EA, Abdu WA, et al: Surgical versus nonoperative treatment for lumbar disc herniation: four-year results for the Spine Patient Outcomes Research Trial (SPORT). **Spine (Phila Pa 1976) 33:**2789–2800, 2008
23. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, et al: Surgical vs nonoperative treatment for lumbar disk herniation. The Spine Patient Outcomes Research Trial (SPORT): a randomized trial. **JAMA 296:**2441–2450, 2006

Manuscript submitted March 13, 2014.
Accepted April 3, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14271.

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.

J Neurosurg Spine 21:54–61, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 9: Lumbar fusion for stenosis with spondylolisthesis

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

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

Patients presenting with stenosis associated with a spondylolisthesis will often describe signs and symptoms consistent with neurogenic claudication, radiculopathy, and/or low-back pain. The primary objective of surgery, when deemed appropriate, is to decompress the neural elements. As a result of the decompression, the inherent instability associated with the spondylolisthesis may progress and lead to further misalignment that results in pain or recurrence of neurological complaints. Under these circumstances, lumbar fusion is considered appropriate to stabilize the spine and prevent delayed deterioration. Since publication of the original guidelines there have been a significant number of studies published that continue to support the utility of lumbar fusion for patients presenting with stenosis and spondylolisthesis. Several recently published trials, including the Spine Patient Outcomes Research Trial, are among the largest prospective randomized investigations of this issue. Despite limitations of study design or execution, these trials have consistently demonstrated superior outcomes when patients undergo surgery, with the majority undergoing some type of lumbar fusion procedure. There is insufficient evidence, however, to recommend a standard approach to achieve a solid arthrodesis. When formulating the most appropriate surgical strategy, it is recommended that an individualized approach be adopted, one that takes into consideration the patient's unique anatomical constraints and desires, as well as surgeon's experience.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14274)

Key Words • **fusion** • **lumbar spine** • **spondylolisthesis** • **stenosis** • **practice guidelines**

Recommendations

There is no evidence that conflicts with the previous recommendations formulated from the first iteration of the Lumbar Fusion Guidelines.

Grade B

Surgical decompression and fusion is recommended as an effective treatment alternative for symptomatic stenosis associated with a degenerative spondylolisthesis in patients who desire surgical treatment.

Although there is insufficient evidence to recommend a standard fusion technique, the patient's anatomy, desires, and concerns as well as surgeon experience should all be factored into the decision-making process when determining the optimal strategy for an individual patient to maximize fusion potential while minimizing risk of complications.

Abbreviations used in this paper: ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; PLIF = posterior lumbar interbody fusion; SPORT = Spine Patient Outcomes Research Trial; TLIF = transforaminal lumbar interbody fusion; VAS = visual analog scale.

Part 9: Lumbar fusion for stenosis with spondylolisthesis

Rationale

Patients presenting with clinically relevant stenosis associated with a spondylolisthesis may report signs and symptoms consistent with neurogenic claudication, radiculopathy, and/or low-back pain. A decompressive procedure is often required to alleviate the symptoms associated with the neurological compression syndrome; however, decompression alone can result in progression of the vertebral misalignment. In the original version of the Lumbar Fusion Guidelines, incorporating a posterolateral lumbar fusion (PLF) as an adjunct to a lumbar decompression was considered an appropriate treatment alternative to prevent deformity progression and improve patient outcomes. Supplementation of the PLF with pedicle screw stabilization was considered an appropriate option in the presence of a kyphosis or if instability was suspected.[26] The purpose of the current Guideline Update was to examine the current literature investigating the role of surgical intervention for patients with symptomatic stenosis associated with spondylolisthesis and focus on the utility of lumbar fusion in this patient population.

Literature Search

Several well-publicized randomized controlled clinical trials have been published since the last systematic review published in 2005.[25] Accordingly, the literature search strategy was designed to reflect the existence of potentially high-quality evidence. The National Library of Medicine and the Cochrane Library were searched for articles published between July 2003 and December 2011, using an electronic literature search engine (PubMed and the Cochrane Search Engine, respectively) with the following subject headings: ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ("Spondylolisthesis"[MeSH] OR spondylolisthesis[title]) AND (("2003"[PDAT]: "3000"[PDAT]) AND "humans" [MeSH Terms] AND English[lang]). A total of 134 references were identified. The titles and abstracts of these 134 references were reviewed. Duplicates were discarded, as were nonsystematic reviews, case series, and retrospective cohort studies with fewer than 100 patients. Studies focused on nuances of technique (i.e., choice of bone graft material for fusion) without comparison with nonoperated or nonfused patients were discarded. Studies comparing substantially different procedures (i.e., interbody vs posterolateral fusion) were included in the literature review. Non–English language references were included if there was sufficient translation of key portions of the reference to allow review. The reference lists of previously published systematic reviews were also reviewed to confirm completeness of the literature search. This strategy resulted in 26 primary references and 5 systematic reviews.[1–25,27–32] Ten papers published since the previous review and one paper that was missed in the previous review providing Level III evidence or better are detailed in the evidentiary table (Table 1).

Scientific Foundation

Surgery Versus No Surgery

Weinstein et al.,[29,30] through publication of the Spine Patient Outcomes Research Trial (SPORT) studies, provide the most powerful evidence supporting the role of surgical intervention in patients with stenosis associated with degenerative spondylolisthesis. This large (> 600 patient) multicenter prospective study was originally designed as a randomized trial, but flaws in the study design and the substantial crossover rate between treatment cohorts have led most, including the authors of this study, to focus on the results of the as-treated analysis. As a result, the randomization process was abandoned and the study regarded as a large well-controlled prospective cohort study. The SPORT group demonstrated that when patients are able to select their treatment strategy based on their symptoms, values, and surgical recommendation, those who choose surgery experience superior outcomes in every clinical measure and at every time point for at least 4 years following treatment. It is important to note that surgeons treated patients with decompression and fusion and were free to offer patients whatever technique of decompression and fusion they thought appropriate.[29,30] As a result of the study limitations, the SPORT provides Level II evidence in support of decompression and fusion for stenosis associated with a spondylolisthesis.

In a companion study, Pearson and the SPORT investigators reviewed preoperative radiographic measurements and 1-year follow-up data in an attempt to identify prognostic indicators of outcome following operative or nonoperative management.[24] Patients in the surgical cohort exhibited superior outcomes compared with those treated nonoperatively; however, there were no preoperative radiographic features that predicted ultimate success. This finding was confounded by the fact that the choice of fusion technique was left to the discretion of the treating surgeons. In the nonoperative arm, better outcomes were paradoxically associated with increased mobility at the level of the listhesis. Confounding factors between the "stable" and "hypermobile" groups such as sex, work status, and compensation status make it difficult to interpret these results. The strength of this study is reduced to Level III evidence supporting the role of surgery for stenosis associated with spondylolisthesis.[24]

Surgical Technique

Abdu et al.[1] reviewed the results from the SPORT lumbar spondylolisthesis study and compared results across fusion techniques. The beneficial effects of surgery were maintained over 4 years, and patients reported significant improvement in every primary outcome measure (Oswestry Disability Index [ODI], 36-Item Short Form Health Survey, and visual analog scale [VAS]) compared with their baseline status. No differences in outcome were detected between the different fusion cohorts (noninstrumented PLF, instrumented PLF, and a 360° approach, instrumented PLF with an interbody graft). The potential for bias exist, however, because surgeons were free to choose the fusion technique, there were impor-

TABLE 1: Lumbar fusion for stenosis with spondylolisthesis: summary of evidence*

Authors & Year	Level of Evidence	Description	Results	Conclusion
Abdu et al., 2009	III: non-controlled cohort study downgraded to III because groups not similar in several important areas.	Four-year follow-up of patients in the surgical arm of SPORT; examined effect of surgical technique on long-term outcome.	All surgical groups exhibited continued improvement in all primary outcomes measures out to 4 yrs. At 4 yrs, there were no differences in outcomes btwn different fusion techniques (noninstrumented, instrumented PLF, instrumented + interbody fusion).	All surgical groups had continued improvement in all primary outcome measures. Surgeons should choose fusion technique based on experience & patient characteristics.
Anderson et al., 2006	I for select population, II for overall spondylolisthesis population.	Two-year follow-up of X-STOP vs nonoperative management for patients w/ mild/moderate claudication & Grade I spondylolisthesis; subgroup analysis of overall RCT.	Patients treated w/ the X-STOP device had better outcomes on Zurich Pain Questionnaire & Patient Satisfaction Scores than patients treated nonoperatively.	Patients w/ mild-moderate claudication due to stenosis associated w/ Grade I spondylolisthesis had better outcomes following X-STOP surgery than those receiving nonoperative treatment.
Cheng et al., 2009	II for functional outcome, III for fusion rate. Downgraded due to inclusion of isthmic spondylolisthesis patients & use of static radiographs for fusion assessment.	RCT w/ 4-yr outcomes comparing PLF w/ PLIF for spondylolisthesis.	Fusion rates higher in the PLIF group, & instrumentation failure higher in the PLF group, but overall functional outcomes the same in both. Note high percentage of isthmic & Grade II spondylolisthesis. Use of interbody techniques may be more advantageous in higher-grade slips & in isthmic deformities.	Good outcomes associated w/ either PLF or PLIF in patients w/ spondylolisthesis. Interbody techniques may have advantages in cases of higher-grade slips or greater instability.
Fernández-Fairen et al., 2007	II	RCT of 82 patients treated w/ PLF w/ either unilat or bilat pedicle screws.	No differences in fusion rates or SF-36 scores between groups. Screw malposition rate, OR time, & blood loss all lower in unilat group.	When performing PLF for low-grade degenerative spondylolisthesis, unilat screws appear to provide similar benefit to bilat screws w/ fewer complications.
Inamdar et al., 2006	II for outcomes (downgraded for small size & short follow-up); III for fusion rates due to reliance on static radiographs.	Very small RCT (20 patients) looking at PLIF vs PLF as adjunct to decompression for stenosis w/ spondylolisthesis.	No significant differences in fusion rates or outcomes btwn patients treated w/ PLF or PLIF when assessed 1 year postop.	Authors recommend PLF as opposed to PLIF due to similar results w/ fewer complications & costs.
Kanayama et al., 2006	III diagnostic (only those felt to be fused were explored).	Very small RCT looking at OP-1 vs autograft & ceramic as fusion substrate in spondylolisthesis population.	CT & dynamic radiographs used to assess fusion, & those patients felt to be fused were explored. Five of the 16 patients thought to be fused based on CT criteria were found to have pseudarthrosis.	Fusion status difficult to establish even w/ CT & dynamic studies.
Kornblum et al., 2004	III (prospective case series or possibly case control).	Prospective series looking at longer-term outcomes (7–8 yrs) in noninstrumented group from Fischgrund et al. study.	Patients w/ solid arthrodesis following noninstrumented fusion based on dynamic radiographs had better functional outcomes than those w/ pseudarthrosis. However, patients w/ pseudarthrosis had greater preop mobility than those who achieved fusion.	Occurrence of a fusion based on dynamic radiographs associated w/ improved functional outcomes. Patients w/ increased preop mobility may benefit from adjuncts that improve fusion rates.
McGuire & Amundson, 1993	III (small size, select population, nonblinded).	Small RCT in select population (mean age 35 yrs, military) evaluating influence of instrumentation on fusion rates following PLF.	No significant difference in fusion rates btwn instrumented & noninstrumented groups at 2-yr follow-up. Functional outcomes assessed w/ nonvalidated instrument.	In young healthy patients, addition of instrumentation to PLF does not appear to improve fusion rates.

(continued)

TABLE 1: Lumbar fusion for stenosis with spondylolisthesis: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Results	Conclusion
Pearson et al., 2008	III (short follow-up & subgroup analysis).	Patients in the SPORT were evaluated in order to assess if there were pretreatment radiographic features that predicted successful treatment.	In the operative arm of the study, there were no particular radiographic features that predicted success or failure at 1 yr. Outcomes superior in the surgical group. Patients w/ greater instability improved more than others in the nonoperative group.	Patient outcomes improve w/ surgery compared to nonoperative treatment when appropriate patients are offered appropriate treatment.
Weinstein et al., 2007	II (crossover resulted in large prospective cohort study).	Randomized (n = 304) & observational (n = 303) trial comparing operative & nonoperative treatment for patients w/ symptomatic stenosis w/ associated spondylolisthesis.	Patients treated surgically had better outcomes on every outcome measure & at every time point up to 2 yrs postop.	Surgical intervention associated w/ superior outcomes compared to nonsurgical measures in patients w/ symptomatic stenosis & spondylolisthesis whose symptoms warrant intervention.
Weinstein et al., 2009	II (crossover resulted in large prospective cohort study).	Four-year follow-up on 2007 study.	Benefits of surgical vs nonsurgical intervention persist at 4-yr follow-up.	Surgery associated w/ superior outcomes than nonoperative measures in patients w/ symptomatic stenosis & spondylolisthesis whose symptoms warrant intervention. This benefit persists through at least 4 yrs.

* OP-1 = osteogenic protein–1; OR = operating room; PLF = posterolateral lumbar fusion; PLIF = posterior lumbar interbody fusion; RCT = randomized controlled trial; SF-36 = 36-Item Short Form Health Survey; SPORT = Spine Patient Outcomes Research Trial.

tant demographic differences between the fusion groups (age and race for example), and there were potential differences not described (such as the degree of disc space collapse or regional kyphosis). These confounding factors limit the ability to formulate relevant conclusions regarding the equivalence or nonequivalence of the various fusion techniques.[1]

Cheng and colleagues[9] performed a randomized trial to evaluate the differences between PLF and posterior lumbar interbody fusion (PLIF) following decompression in a group of 138 patients with degenerative or isthmic spondylolisthesis (Grade I or II). They found that fusion rates were higher and instrumentation-related complication rates were lower in the PLIF group. However, functional outcomes were identical between the groups, and the study relied on static radiographs for the assessment of fusion. The fact that the majority of patients had isthmic spondylolisthesis and that a high percentage of patients had Grade II slips decreases the generalizability of these data to the degenerative population. Due to the heterogeneous patient population and questionable criteria to assess fusion status, the study was downgraded to Level II evidence in support of a PLF or PLIF following decompression for the treatment of degenerative spondylolisthesis. Consideration of interbody techniques may be appropriate in patients with higher-grade slips.[9]

Fernández-Fairen and colleagues[12] performed a randomized trial in a cohort of 82 patients in whom they examined the effect of unilateral versus bilateral screw fixation as an adjunct to PLF following decompression for degenerative spondylolisthesis. While the sample size was relatively small, the study was powered to detect significant differences on validated outcomes measures and CT scanning was used to determine fusion status 3 years after surgery. The authors group observed no differences in functional outcomes or in fusion rates between the 2 groups and found that complication rates, blood loss, and operative time were lower in the group in which unilateral screws were placed. This study provides Level II evidence that unilateral screw fixation is associated with similar outcomes as bilateral screw fixation, but because the data are generated from a single study with a relatively small patient population, the validity of this conclusion is limited.

Inamdar et al.[16] performed a randomized study involving 20 patients to investigate the differences in outcomes between PLF and PLIF following decompression for stenosis associated with spondylolisthesis. Clinical and radiographic follow-up data were limited to 1 year. Fusion status was assessed using static radiographs. Although no differences were detected between the treatment groups, the small sample size, short follow-up duration, and questionable method of fusion assessment compromise the conclusions formulated by the authors; therefore, this study is downgraded to Level II evidence in support of PLF over PLIF (Level II for outcomes and Level III for fusion status).[16]

Kornblum and colleagues[19] followed up the noninstrumented cohort from the Fischgrund et al. study[13] for a mean of 7.7 years. They followed up 47 of the original 58 patients: only 1 patient was lost to follow-up, 8 died, 1

was disabled from a stroke, and 1 declined to participate. They found that patients in this group who were thought to have a solid arthrodesis (based on dynamic radiographs) enjoyed better functional outcomes (as measured using VAS for pain assessment and the Stucki inventory) than patients treated with the same procedure in whom a solid arthrodesis was not achieved.[13,19] It was noted that those patients in whom arthrodesis was not achieved had significantly greater preoperative angular mobility. This paper provides Level III evidence as a case-control study showing that efforts to increase fusion rates are associated with better outcomes in patients treated with fusion as an adjunct to decompression.

McGuire and Amundson[20] studied a military population of patients with stenosis and spondylolisthesis and randomized a total of 27 patients to decompression and fusion with or without instrumentation. Fusion rates at 2 years, based on assessment of flexion-extension radiographs, were similar between the groups (72% without instrumentation vs 78% with instrumentation). This paper is felt to provide Level III evidence (small study, nonblinded, very select population with mean age of 35 years) that the addition of instrumentation does not improve fusion rates.[20] This paper was not included in the previous systematic review.[25]

Other papers have been discussed previously or provide lower-quality evidence. Since some of these provided the basis for the past recommendations, they are briefly discussed below.

Andersen et al.[2] described long-term outcomes following instrumented and noninstrumented fusion for chronic low-back pain but did not separate out patients with degenerative lumbar spondylolisthesis. This is the same patient cohort previously described by Bjarke Christensen et al.[6]

Athiviraham and Yen[5] described a cohort series of patients treated nonoperatively, with decompression alone, or with decompression and fusion. Only patients with spondylolisthesis underwent fusion. Due to this important difference between the patient groups in this prospective comparison, this paper is felt to provide only Level IV evidence.

Bridwell and colleagues[7] performed a pseudo-randomized study involving 43 patients treated operatively for stenosis associated with spondylolisthesis. Nine patients underwent decompression alone; 10, decompression and noninstrumented PLF; and 24, decompression and instrumented PLF. Functional outcomes were better in the fusion group, and better functional outcomes were associated with arrest of slip progression and solid fusion. The use of instrumentation appeared to improve fusion rates as well as patient outcomes. The study was downgraded to a Level III study because the investigators used nonvalidated outcomes measures and relied on static radiographs for the determination of fusion.[7] This paper was previously reviewed in the 2005 Fusion Guidelines.[25]

Carreon and colleagues[8] performed a systemic review of the literature to evaluate the effects of fusion on different patient populations. They found that the presence of an established diagnosis such as spondylolisthesis was associated with better functional outcomes compared with patients treated with similar procedures for chronic low-back pain without a demonstrable deformity. Because the analysis included very few spondylolisthesis patients (96 of 2002) and because the index studies are discussed elsewhere in this Guideline Update, the Carreon et al. review does not provide unique information regarding the treatment of this patient population. It does provide supporting evidence confirming that good outcomes may be expected in patients treated with fusion for degenerative spondylolisthesis.

Chou et al.[10] performed a systematic review of the literature regarding the surgical versus nonsurgical management of low-back pain. While fusion for patients with stenosis was evaluated, spondylolisthesis and nonspondylolisthesis groups were considered together. No specific information regarding the treatment of patients with stenosis and associated spondylolisthesis is given.

Christensen and colleagues[11] randomized 130 patients with isthmic spondylolisthesis, primary degenerative instability (back pain associated with movement and degenerative disc disease), or secondary degenerative instability (same as primary but with history of having undergone decompression) to PLF with or without instrumentation. No differences between the 2 groups were detected; however, the patient population is not relevant to a discussion of patients with stenosis and degenerative spondylolisthesis. Andersen et al.[2] described long-term outcomes following instrumented and noninstrumented fusion for chronic low-back pain but did not separate out patients with degenerative lumbar spondylolisthesis. This is the same patient cohort previously described by Bjarke Christensen et al.[6]

Fischgrund and colleagues[13] performed a prospective clinical trial of 68 patients with stenosis and degenerative spondylolisthesis who were randomized into one of 2 groups: decompression and PLF in one group and decompression and PLF supplemented with pedicle screw fixation in the other. Fusion status was assessed using plain and dynamic radiography, and clinical outcomes were assessed using a VAS for pain as well as a patient satisfaction scale. The patients treated with pedicle screw fixation had a statistically significantly higher fusion rate (83%) than those treated with noninstrumented fusion (45%). Both groups demonstrated significant score improvements on the VAS for both back and leg pain ($p = 0.001$), and the majority of patients in both groups reported their outcomes as good or excellent (78% in the instrumented group and 85% in the noninstrumented group). This paper provides Level I medical evidence that pedicle screw fixation, as an adjunct to decompression and PLF, improves fusion success, and Level III medical evidence (due to the nonvalidated patient satisfaction scale and inadequate sample size), suggesting that pedicle screw fixation does not improve functional outcome following PLF in this patient population.[13] This paper was previously discussed in the 2005 Fusion Guidelines.[25]

Gibson and Waddell[14] performed a systematic review of randomized trials for the Cochrane Review in 2005. The authors did not review any references not reviewed in the previous guidelines document and did not consider patients with stenosis and spondylolisthesis separately.[25]

Kanayama and colleagues[17] performed a small randomized controlled trial comparing osteogenic protein-1 (OP-1) to autograft plus ceramic as fusion materials in a group of 19 patients undergoing instrumented PLF following decompression for stenosis associated with spondylolisthesis. The OP-1 group was found to have a slightly lower fusion rate as judged by CT scans, dynamic radiographs, and exploration. While new bone formation was noted in both groups, patients who underwent surgical reexploration for planned instrumentation removal were found to have a relatively high incidence of nonunion despite CT- and dynamic radiography–documented evidence of fusion. This paper does not contribute much to the discussion of treatment options for patients with stenosis and spondylolisthesis but does provide information regarding the limitations of imaging studies to provide information regarding the presence or absence of fusion (Level III diagnostic study as patients without radiographic fusion were not surgically explored to confirm/refute fusion status).

Kondrashov and colleagues[18] followed up 18 patients treated with the X-STOP device and found that beneficial effects appeared to be durable for a mean of 4.2 years of follow-up in their series (Level IV evidence).

McNeely et al.[21] performed a systematic review of the effect of physiotherapy on back pain in patients with various diagnoses including spondylolisthesis. They found that there was a paucity of evidence to support the effectiveness of physiotherapy for patients with degenerative spondylolisthesis. This paucity is the result of very few studies and the fact that patients with degenerative spondylolisthesis were not necessarily considered separately. Two randomized studies were reviewed: one on younger patients with isthmic spondylolisthesis[23] and the other on patients with chronic low-back pain and a variety of spinal alignments but without claudication.[27]

Mirza and Deyo[22] performed a systematic review of trials evaluating the surgical management of low-back pain. The review did not separately consider patients with stenosis and spondylolisthesis.

Thomsen et al.[28] performed a randomized controlled clinical trial of 130 patients who underwent lumbar fusion for low-back pain. The patients were randomized to instrumented (pedicle screw fixation) and noninstrumented PLF groups. Overall, there was no significant difference in functional outcome (as measured by the Dallas Pain Questionnaire). Although this paper describes a randomized controlled trial with validated outcome measures, the overall patient population was not that of stenosis and associated spondylolisthesis (isthmic spondylolisthesis, primary and secondary degenerative instability). Only a small subgroup of patients underwent decompression, and it is unclear whether these patients had associated spondylolisthesis. This paper was previously reviewed in the 2005 Fusion Guidelines.

Welch et al.[31] provided information regarding a prospective case series of patients with stenosis and degenerative spondylolisthesis who were treated with a dynamic fixation device. Overall results appeared promising; however, no comparison cohort was described. This paper is felt to provide Level IV information regarding the potential utility of dynamic fixation in select patients with stenosis and degenerative spondylolisthesis.[31]

Zucherman et al.[32] performed a prospective randomized study to assess the efficacy of the X-STOP device for the treatment of mild to moderate neurogenic claudication. The results relevant to this discussion have been presented by Anderson et al.[3] and discussed previously.

Summary

The current medical evidence continues to support the role of surgery over nonoperative therapies for patients with symptomatic stenosis associated with spondylolisthesis. The vast majority of patients across these studies underwent an instrumented PLF. The achievement of a solid arthrodesis is associated with superior outcomes, and therefore, efforts to maximize fusion potential should be considered. A variety of surgical alternatives may be considered. Surgeons should choose the technique based on their own experience, the risk of complications, and the individual patient's anatomical and physiological characteristics, comorbidities, and preference. It is recognized, however, that within this patient population significant heterogeneity exists that may have an impact on treatment response.

Key Issues for Future Investigation

The utility of surgical intervention in this patient population is well established. Future work should focus on identifying prognostic indicators of surgical outcome and stratify these factors among the various fusion techniques. Establishing well-designed randomized control trials to address these issues will be extremely difficult if not impractical (as exemplified by the SPORT), but relevant data may be obtained by establishing a prospective diagnosis-based registry.

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., 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 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: Resnick. 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: Resnick. Study supervision: Kaiser.

References

1. Abdu WA, Lurie JD, Spratt KF, Tosteson AN, Zhao W, Tosteson TD, et al: Degenerative spondylolisthesis: does fusion method influence outcome? Four-year results of the spine patient outcomes research trial. **Spine (Phila Pa 1976) 34:**2351–2360, 2009
2. Andersen T, Videbaek TS, Hansen ES, Bünger C, Christensen FB: The positive effect of posterolateral lumbar spinal fusion is preserved at long-term follow-up: a RCT with 11-13 year follow-up. **Eur Spine J 17:**272–280, 2008
3. Anderson PA, Tribus CB, Kitchel SH: Treatment of neurogenic claudication by interspinous decompression: application of the X STOP device in patients with lumbar degenerative spondylolisthesis. **J Neurosurg Spine 4:**463–471, 2006
4. Arts MP, Verstegen MJ, Brand R, Koes BW, van den Akker ME, Peul WC: Cost-effectiveness of decompression according to Gill versus instrumented spondylodesis in the treatment of sciatica due to low grade spondylolytic spondylolisthesis: a prospective randomised controlled trial [NTR1300]. **BMC Musculoskelet Disord 9:**128, 2008
5. Athiviraham A, Yen D: Is spinal stenosis better treated surgically or nonsurgically? **Clin Orthop Relat Res 458:**90–93, 2007
6. Bjarke Christensen F, Stender Hansen E, Laursen M, Thomsen K, Bünger CE: Long-term functional outcome of pedicle screw instrumentation as a support for posterolateral spinal fusion: randomized clinical study with a 5-year follow-up. **Spine (Phila Pa 1976) 27:**1269–1277, 2002
7. Bridwell KH, Sedgewick TA, O'Brien MF, Lenke LG, Baldus C: The role of fusion and instrumentation in the treatment of degenerative spondylolisthesis with spinal stenosis. **J Spinal Disord 6:**461–472, 1993
8. 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
9. Cheng L, Nie L, Zhang L: Posterior lumbar interbody fusion versus posterolateral fusion in spondylolisthesis: a prospective controlled study in the Han nationality. **Int Orthop 33:** 1043–1047, 2009
10. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD: Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. **Spine (Phila Pa 1976) 34:**1094–1109, 2009
11. Christensen FB, Thomsen K, Eiskjaer SP, Hansen ES, Fruensgaard S, Gelinick J, et al: [The effect of pedicle screw instrumentation on posterolateral spinal fusion. A prospective, randomized study with a two-year follow-up.] **Ugeskr Laeger 161:**1920–1925, 1999 (Danish)
12. Fernández-Fairen M, Sala P, Ramírez H, Gil J: A prospective randomized study of unilateral versus bilateral instrumented posterolateral lumbar fusion in degenerative spondylolisthesis. **Spine (Phila Pa 1976) 32:**395–401, 2007
13. Fischgrund JS, Mackay M, Herkowitz HN, Brower R, Montgomery DM, Kurz LT: 1997 Volvo Award Winner in Clinical Studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. **Spine (Phila Pa 1976) 22:**2807–2812, 1997
14. Gibson JN, Waddell G: Surgery for degenerative lumbar spondylosis: updated Cochrane Review. **Spine (Phila Pa 1976) 30:** 2312–2320, 2005
15. Grob D, Humke T, Dvorak J: [Significance of simultaneous fusion and surgical decompression in lumbar spinal stenosis.] **Orthopade 22:**243–249, 1993 (Ger)
16. Inamdar DN, Alagappan M, Shyam L, Devadoss S, Devadoss A: Posterior lumbar interbody fusion versus intertransverse fusion in the treatment of lumbar spondylolisthesis. **J Orthop Surg (Hong Kong) 14:**21–26, 2006
17. Kanayama M, Hashimoto T, Shigenobu K, Yamane S, Bauer TW, Togawa D: A prospective randomized study of posterolateral lumbar fusion using osteogenic protein-1 (OP-1) versus local autograft with ceramic bone substitute: emphasis of surgical exploration and histologic assessment. **Spine (Phila Pa 1976) 31:**1067–1074, 2006
18. Kondrashov DG, Hannibal M, Hsu KY, Zucherman JF: Interspinous process decompression with the X-STOP device for lumbar spinal stenosis: a 4-year follow-up study. **J Spinal Disord Tech 19:**323–327, 2006
19. Kornblum MB, Fischgrund JS, Herkowitz HN, Abraham DA, Berkower DL, Ditkoff JS: Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective long-term study comparing fusion and pseudarthrosis. **Spine (Phila Pa 1976) 29:**726–734, 2004
20. McGuire RA, Amundson GM: The use of primary internal fixation in spondylolisthesis. **Spine (Phila Pa 1976) 18:**1662–1672, 1993
21. McNeely ML, Torrance G, Magee DJ: A systematic review of physiotherapy for spondylolysis and spondylolisthesis. **Man Ther 8:**80–91, 2003
22. Mirza SK, Deyo RA: Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. **Spine (Phila Pa 1976) 32:**816–823, 2007
23. O'Sullivan PB, Phyty GD, Twomey LT, Allison GT: Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. **Spine (Phila Pa 1976) 22:**2959–2967, 1997
24. Pearson AM, Lurie JD, Blood EA, Frymoyer JW, Braeutigam H, An H, et al: Spine Patient Outcomes Research Trial: radiographic predictors of clinical outcomes after operative or nonoperative treatment of degenerative spondylolisthesis. **Spine (Phila Pa 1976) 33:**2759–2766, 2008
25. 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 9: fusion in patients with stenosis and spondylolisthesis. **J Neurosurg Spine 2:**679–685, 2005
26. 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 12: pedicle screw fixation as an adjunct to posterolateral fusion for low-back pain. **J Neurosurg Spine 2:**700–706, 2005
27. Spratt KF, Weinstein JN, Lehmann TR, Woody J, Sayre H: Efficacy of flexion and extension treatments incorporating braces for low-back pain patients with retrodisplacement, spondylolisthesis, or normal sagittal translation. **Spine (Phila Pa 1976) 18:**1839–1849, 1993
28. Thomsen K, Christensen FB, Eiskjaer SP, Hansen ES, Fruensgaard S, Bünger CE: 1997 Volvo Award Winner in Clini-

cal Studies. The effect of pedicle screw instrumentation on functional outcome and fusion rates in posterolateral lumbar spinal fusion: a prospective, randomized clinical study. **Spine (Phila Pa 1976) 22:**2813–2822, 1997
29. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al: Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. **N Engl J Med 356:** 2257–2270, 2007
30. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, et al: Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. Four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. **J Bone Joint Surg Am 91:**1295–1304, 2009
31. Welch WC, Cheng BC, Awad TE, Davis R, Maxwell JH, Delamarter R, et al: Clinical outcomes of the Dynesys dynamic neutralization system: 1-year preliminary results. **Neurosurg Focus 22(1):**E8, 2007
32. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, et al: A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. **Spine (Phila Pa 1976) 30:**1351–1358, 2005

Manuscript submitted March 13, 2014.
Accepted April 3, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14274.
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.

J Neurosurg Spine 21:62–66, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 10: Lumbar fusion for stenosis without spondylolisthesis

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

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

Lumbar stenosis is one of the more common radiographic manifestations of the aging process, leading to narrowing of the spinal canal and foramen. When stenosis is clinically relevant, patients often describe activity-related low-back or lower-extremity pain, known as neurogenic claudication. For those patients who do not improve with conservative care, surgery is considered an appropriate treatment alternative. The primary objective of surgery is to reconstitute the spinal canal. The role of fusion, in the absence of a degenerative deformity, is uncertain. The previous guideline recommended against the inclusion of lumbar fusion in the absence of spinal instability or a likelihood of iatrogenic instability. Since the publication of the original guidelines, numerous studies have demonstrated the role of surgical decompression in this patient population; however, few have investigated the utility of fusion in patients without underlying instability. The majority of studies contain a heterogeneous cohort of subjects, often combining patients with and without spondylolisthesis who received various surgical interventions, limiting fusions to those patients with instability. It is difficult if not impossible, therefore, to formulate valid conclusions regarding the utility of fusion for patients with uncomplicated stenosis. Lower-level evidence exists, however, that does not demonstrate an added benefit of fusion for these patients; therefore, in the absence of deformity or instability, the inclusion of a fusion is not recommended.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14275)

Key Words • **stenosis** • **lumbar spine** • **neurogenic claudication** • **fusion** • **practice guidelines**

Recommendations

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

Grade B

Surgical decompression is recommended for patients with symptomatic neurogenic claudication due to lumbar stenosis without spondylolisthesis who elect to undergo surgical intervention (Level II/III evidence).

Grade C

In the absence of deformity or instability, lumbar fusion has not been shown to improve outcomes in patients with isolated stenosis, and therefore it is not recommended (Level IV evidence).

Rationale

Lumbar stenosis, narrowing of the spinal canal as a

Abbreviations used in this paper: PLF = posterolateral lumbar fusion; ODI = Oswestry Disability Index; VAS = visual analog scale.

consequence of degenerative disease, is a common phenomenon associated with the natural process of aging that can lead to the clinical syndrome known as neurogenic claudication. Patients typically describe activity-related low-back and leg pain that worsens with prolonged standing or ambulation, compromising their quality of life. This is a relatively common disorder, particularly among the elderly, that can lead to significant disability. In the absence of an associated spinal deformity or instability, symptoms of neurogenic claudication typically respond to decompression in patients whose presentation and general health warrant operative intervention. The inclusion of lumbar fusion in the surgical management of this patient population is unclear.

Literature Search

The National Library of Medicine was searched from July 2003 to December 2011 using the Internet-based search engine PubMed with the following search terms: ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ("Spinal Stenosis"[MeSH] OR stenosis[title]). The search was limited to the English language and human subjects. A total of 174 references were retrieved. The Cochrane database was also searched using the same search terms, and no additional references were identified. The titles and abstracts of these references were reviewed, and papers dealing with basic science or patients presenting with spondylolisthesis or degenerative scoliosis were excluded, as were case reports, editorials, and nonstructured reviews. Thirty-six references were identified that provide either background information or new data regarding the role of fusion in patients with stenosis without spondylolisthesis or scoliosis. Studies providing comparative data between fusion and nonfusion procedures serve as the scientific foundation of this review and are summarized in Table 1.

Scientific Foundation

The benefits of surgical decompression for lumbar stenosis, coupled with a fusion in the presence of radiographic instability or spondylolisthesis, have been well documented. Malmivaara et al. conducted a randomized clinical trial evaluating the effectiveness of surgical versus nonsurgical intervention in 94 patients with mild to moderate symptoms of neurogenic claudication due to spinal stenosis.[8] Although the investigators performed a power analysis to determine sample size, the number of patients included in each cohort did not meet the predetermined threshold. Objective validated outcome instruments were used to assess clinical status prior to surgery and at 6, 12, and 24 months after surgery. At the discretion of the treating surgeon, 10 of the 50 patients in the surgical cohort underwent fusion, with or without instrumentation, because of the presence of spondylolisthesis. Patients who were treated surgically had statistically significant clinical improvements in the Oswestry Disability Index (ODI) and visual analog scale (VAS) compared with those treated nonoperatively, which persisted over the study period. Study limitations, however, did exist, including small sample size, heterogeneity of the patient population and surgical intervention, and unblinded assessment of clinical and radiographic outcome. Due to these limitations, the paper was downgraded to Level II evidence, supporting surgical decompression as an effective modality for patients with mild to moderate symptoms of neurogenic claudication due to lumbar stenosis. Because of the small number of patients undergoing fusion, no valid comparison between decompression and decompression with fusion can be performed.

Athiviraham and Yen performed a prospective cohort study in a group of 125 patients comparing operative to nonoperative management for neurogenic claudication due to lumbar stenosis.[1] Patients with isolated stenosis underwent lumbar decompression, while those with an associated spondylolisthesis underwent fusion. Overall outcomes were substantially improved in both surgical groups compared with the nonsurgical cohort. Due to several study design limitations, including small sample size and potential for selection bias, this investigation provides Level III evidence in support of operative intervention for the treatment of spinal stenosis.

Despite the evidence supporting the utility of lumbar fusion for patients presenting with spondylolisthesis or radiographic instability, there remains considerable debate with respect to patients presenting only with stenosis. Although clinical success has been documented in patients undergoing both decompression and fusion for stenosis, the majority of these studies are based on a compromised study design and provide low levels of evidence. Grivas et al. performed a retrospective review of 23 patients who were treated with decompression and fusion with or without instrumentation for neurogenic claudication due to lumbar stenosis.[4] Five of the 23 patients had an associated spondylolisthesis. The authors found that all patients showed improvements on the 36-Item Short Form Health Survey, with the instrumented group showing a greater improvement. This case series provides Level IV evidence that improved outcomes can be achieved with decompression and fusion in the lumbar stenosis population, but it does not provide any evidence regarding the relative benefit of fusion in addition to decompression.

Gu et al. performed a retrospective review of 81 patients who underwent surgery for neurogenic claudication due to lumbar stenosis.[5] Forty-three patients were treated with decompression and posterolateral lumbar fusion (PLF), and 38 were treated with decompression and instrumented PLF. All patients were subsequently treated with 3–4 weeks of bed rest followed by gradual mobilization. Both groups of patients improved, and there were no differences in outcomes between the groups at a mean follow-up of 6.2 years. Both groups included a fair number of patients with spondylolisthesis or radiographic evidence of instability. While there were no overall differences between the groups with regard to the presence or absence of spondylolisthesis or radiographic instability preoperatively, the authors stated that they preferred to use instrumentation in younger or more active patients. The overall success rate was just over 70% in both groups.

TABLE 1: Lumbar fusion for stenosis without spondylolisthesis: summary of evidence*

Authors & Year	Description	Level of Evidence	Results	Conclusion
Malmivaara et al., 2007	Randomized controlled trial comparing surgical intervention w/ nonsurgical intervention in a group of pts w/ moderate symptoms of neurogenic claudication. A power analysis based on a presumed 15-point difference on the ODI was performed, as was an intent-to-treat analysis.	II for decompression vs nonop intervention in pts w/ moderate symptoms.	94 pts randomized w/ minimal crossover & loss to follow-up. Pts treated w/ decompression did statistically & clinically significantly better than those treated nonoperatively. Fusion was reserved for 10 pts w/ concomitant spondylolisthesis.	Decompression is an effective treatment for neurogenic claudication due to LSS. Fusion is appropriate in pts w/ coexisting spondylolisthesis.
Trouillier et al., 2004	Retrospective series of 85 pts undergoing surgery for LSS & followed for a mean of 79 mos. Pts were treated w/ minimal decompression, extensive decompression, or decompression & instrumented fusion depending on radiographic findings.	IV: Retrospective series of pts treated w/ decompression w/ or w/o fusion.	79 of 85 pts were followed, & all 3 groups exhibited improved symptoms. Those treated w/ more minimal surgical procedures tended to do better; however, there was a significant incidence of late instability in the extensive decompression w/o fusion group.	Decompression is effective for relieving symptoms of neurogenic claudication. Fusion is appropriate in cases where there is preop or intraop evidence of instability.

* LSS = lumbar spinal stenosis; pts = patients.

This paper does not address the issue of fusion versus no fusion in the lumbar stenosis population without deformity or instability.

Jansson et al. performed a retrospective review of 9664 operations performed for lumbar stenosis in the Swedish population with 10-year follow-up and reported a reoperation rate of 11%.[6,7] Eighty-nine percent of patients were treated with laminectomy alone, and 11% were treated with laminectomy and fusion with or without instrumentation. They noted that reoperation rates were lower in patients who had undergone a fusion in addition to decompression as opposed to decompression alone. Because the data were drawn from an administrative database and because no information is provided regarding why the patients were selected for fusion versus nonfusion procedures, the study does not provide useful information with regard to the benefit of fusion as an adjunct to decompression for lumbar stenosis without deformity or instability.

Rampersaud et al. performed an interesting study comparing benefits measured by standard health utility indexes between patients treated with surgery (decompression with or without fusion) for lumbar spinal stenosis or with joint arthroplasty of the hip and knee.[9] The authors found that benefits were comparable or superior in the group treated for lumbar stenosis over 2 years. Patients with spondylolisthesis were included in the stenosis group. While no differences were detected between the fusion and nonfusion subgroups of patients who underwent surgery for lumbar stenosis, differences in selection criteria make a direct comparison impossible. This paper provides Level IV evidence regarding the relative effectiveness of surgery for neurogenic claudication due to lumbar stenosis but does not provide useful information regarding the utility of fusion in patients without deformity.

Trouillier and colleagues performed a retrospective review of 85 patients who underwent surgery for neurogenic claudication.[11] Patients were treated with minimal decompression, extensive decompression, or extensive decompression and instrumented fusion. Surgical decision making was dependent on the severity of stenosis, as determined by preoperative myelography, and/or the presence of instability, defined either on preoperative imaging or during intraoperative assessment. Patient response was measured utilizing validated outcome measures, including the ODI and VAS, for a mean follow-up period of 79 months. All patients improved; however, patients with less extensive surgery tended to do better. Six of 16 patients with extensive decompressions without fusion developed radiographic evidence of instability, defined as greater than 5-mm translation on dynamic radiographs. The paper provides Level IV evidence on the effectiveness of decompression for symptoms of neurogenic claudication and supports the role of fusion in cases in which there is preoperative radiographic or intraoperative evidence of iatrogenic instability.

Yamashita et al. studied the relationship between functional disability, patient satisfaction, and walking ability in a cohort of 77 patients who were treated with decompression with or without fusion.[13] They found that patients improved in all outcomes measures but that patient satisfaction was not always tied to functional improvement as defined by the ODI. Persistent difficulty in walking was associated with lower patient satisfaction. Patients were chosen for fusion based on the preoperative diagnosis of spondylolisthesis, so no comparison between decompression alone or decompression plus fusion can be made.

Zouboulis and colleagues performed a prospective evaluation of a group of 41 patients who were treated with laminectomy and instrumented fusion for stenosis.[14] The

patient group was mixed and contained patients with normal alignment, scoliosis, spondylolisthesis, and multilevel disease. Overall, functional outcomes were improved over a mean of 3.7 years. Ninety-five percent of patients reported satisfactory results; however, 3 patients required further stabilization surgery during the follow-up period. This paper provides Level IV evidence that decompression and fusion provides benefit to some patients with lumbar stenosis, but it does not provide useful evidence regarding the role of fusion in patients without deformity or instability.[14]

Previous structured reviews of the literature have been performed using a variety of methodologies. Resnick et al., Gibson and Waddell, Watters et al., and Chou et al. all reviewed the available literature and all concluded that in the absence of deformity or instability, the performance of lumbar fusion was not associated with improved outcomes compared with decompression alone.[2,3,10,12]

Summary

Recent publications continue to support the role of surgical intervention over nonoperative management strategies for the treatment of symptomatic lumbar stenosis. For those patients presenting with uncomplicated lumbar stenosis, the literature has consistently demonstrated a beneficial role of lumbar decompression.

To date, there have been no high-quality studies comparing the efficacy of simple decompression with decompression and fusion in patients presenting with stenosis without an associated degenerative deformity. The majority of studies are compromised by a heterogeneous cohort of patients with respect to presenting diagnosis and a lack of standardized surgical approaches. Formulating valid conclusions comparing decompression with decompression and fusion is therefore impossible. In fact, the true effect of lumbar fusion for uncomplicated stenosis cannot be determined since most, if not all, of these studies reserve lumbar fusion for those patients presenting with stenosis and an associated spondylolisthesis.

Key Issues for Future Investigation

It seems highly unlikely that a well-designed investigation will be conducted or is required to compare the efficacy of lumbar decompression with decompression and fusion in patients presenting with uncomplicated lumbar stenosis. It seems more plausible that creation of prospective patient registries will allow the identification of a specific subgroup of patients presenting with routine lumbar stenosis that may benefit from the inclusion of a lumbar fusion. Once this profile is established, a more comprehensive well-designed comparative study could be conducted to determine the true treatment effect of lumbar 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 Cockroft, 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 Pilitsis, M.D., Ph.D.; Joshua Rosenow, M.D., F.A.C.S.; 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: Resnick. 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: Resnick. Study supervision: Kaiser.

References

1. Athiviraham A, Yen D: Is spinal stenosis better treated surgically or nonsurgically? **Clin Orthop Relat Res 458:**90–93, 2007
2. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD: Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. **Spine (Phila Pa 1976) 34:**1094–1109, 2009
3. Gibson JN, Waddell G: Surgery for degenerative lumbar spondylosis: updated Cochrane Review. **Spine (Phila Pa 1976) 30:**2312–2320, 2005
4. Grivas TB, Vasiliadis E, Papadakis SA, Mouzakis V, Segos D: Quality of life after surgical decompression of lumbar spinal stenosis with and without instrumentation. **Stud Health Technol Inform 123:**456–460, 2006
5. Gu Y, Chen L, Yang HL, Chen XQ, Dong RB, Han GS, et al: Efficacy of surgery and type of fusion in patients with degenerative lumbar spinal stenosis. **J Clin Neurosci 16:**1291–1295, 2009
6. Jansson KA, Blomqvist P, Granath F, Németh G: Spinal stenosis surgery in Sweden 1987–1999. **Eur Spine J 12:**535–541, 2003
7. Jansson KA, Németh G, Granath F, Blomqvist P: Spinal stenosis re-operation rate in Sweden is 11% at 10 years—a national analysis of 9,664 operations. **Eur Spine J 14:**659–663, 2005
8. Malmivaara A, Slätis P, Heliövaara M, Sainio P, Kinnunen H, Kankare J, et al: Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. **Spine (Phila Pa 1976) 32:**1–8, 2007

9. Rampersaud YR, Ravi B, Lewis SJ, Stas V, Barron R, Davey R, et al: Assessment of health-related quality of life after surgical treatment of focal symptomatic spinal stenosis compared with osteoarthritis of the hip or knee. **Spine J 8:**296–304, 2008
10. 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 10: fusion following decompression in patients with stenosis without spondylolisthesis. **J Neurosurg Spine 2:**686–691, 2005
11. Trouillier H, Birkenmaier C, Kluzik J, Kauschke T, Refior HJ: Operative treatment for degenerative lumbar spinal canal stenosis. **Acta Orthop Belg 70:**337–343, 2004
12. Watters WC III, Baisden J, Gilbert TJ, Kreiner S, Resnick DK, Bono CM, et al: Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. **Spine J 8:**305–310, 2008
13. Yamashita K, Ohzono K, Hiroshima K: Five-year outcomes of surgical treatment for degenerative lumbar spinal stenosis: a prospective observational study of symptom severity at standard intervals after surgery. **Spine (Phila Pa 1976) 31:**1484–1490, 2006
14. Zouboulis P, Karageorgos A, Dimakopoulos P, Tyllianakis M, Matzaroglou C, Lambiris E: Functional outcome of surgical treatment for multilevel lumbar spinal stenosis. **Acta Orthop 77:**670–676, 2006 (Erratum in **Acta Orthop 78:**862, 2007)

Manuscript submitted March 13, 2014.
Accepted April 8, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14275.

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.

J Neurosurg Spine 21:67–74, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 11: Interbody techniques for lumbar fusion

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

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

Interbody fusion techniques have been promoted as an adjunct to lumbar fusion procedures in an effort to enhance fusion rates and potentially improve clinical outcome. The medical evidence continues to suggest that interbody techniques are associated with higher fusion rates compared with posterolateral lumbar fusion (PLF) in patients with degenerative spondylolisthesis who demonstrate preoperative instability. There is no conclusive evidence demonstrating improved clinical or radiographic outcomes based on the different interbody fusion techniques. The addition of a PLF when posterior or anterior interbody lumbar fusion is performed remains an option, although due to increased cost and complications, it is not recommended. No substantial clinical benefit has been demonstrated when a PLF is included with an interbody fusion. For lumbar degenerative disc disease without instability, there is moderate evidence that the standalone anterior lumbar interbody fusion (ALIF) has better clinical outcomes than the ALIF plus instrumented, open PLF. With regard to type of interbody spacer used, frozen allograft is associated with lower pseudarthrosis rates compared with freeze-dried allograft; however, this was not associated with a difference in clinical outcome.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14276)

Key Words • **fusion** • **lumbar spine** • **bone graft** • **spondylosis** • **practice guidelines**

Recommendations

There is no evidence that conflicts with the previous recommendations formulated from the first generation of Lumbar Fusion Guidelines published in the original version of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine."

Grade B

The addition of an interbody fusion is recommended as an option to enhance the fusion rate (which lowers the reoperation rate) in patients undergoing lumbar fusion. However, the improvement in fusion rates with the addition of interbody fusion has not consistently translated to an improvement in clinical outcomes (multiple Level II reports).

The addition of posterolateral lumbar fusion (PLF) to interbody fusion is not recommended in patients undergoing lumbar interbody fusion since the evidence indi-

Abbreviations used in this paper: ALIF = anterior lumbar interbody fusion; DDD = degenerative disc disease; FRA = femoral ring allograft; LOS = length of stay; ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; PLIF = posterior lumbar interbody fusion; PPS = instrumented PLF with pedicle screws; SF-36 = 36-Item Short Form Health Survey; TLIF = transforaminal lumbar interbody fusion; VAS = visual analog scale.

cates no substantial clinical benefit but an increased rate of complications if a PLF is added to an interbody fusion (Level II and III reports).

Grade C

Anterior lumbar interbody fusion (ALIF) performed with a frozen femoral ring allograft (FRA) has a lower pseudarthrosis rate than ALIF performed with a freeze-dried FRA for the treatment of degenerative disc disease with or without spondylolisthesis. However, the improved fusion rate did not affect clinical outcomes (Level II evidence from a single report).

Anterior lumbar interbody fusion has better clinical outcomes and fewer perioperative morbidities than instrumented PLF, although the fusion rate is similar between the 2 techniques (Level III evidence from 2 reports).

Rationale

The surgical treatment of degenerative disease of the lumbar spine has evolved over the last several decades, and interbody techniques have been proposed as surgical alternatives to supplement or replace PLF. Placement of the graft within the load-bearing column of the spine has biomechanical advantages and has been reported to result in higher fusion rates with improved patient outcomes compared with PLF techniques. A variety of techniques are available for the application of interbody grafts, and each technique has particular advantages and disadvantages. The purpose of this review is to examine the current evidence investigating the experience with interbody fusion techniques and their relative safety and efficacy compared with PLF techniques for the treatment of patients with degenerative lumbar disease.

Literature Search

A computerized search of the National Library of Medicine MEDLINE database, utilizing the online search engine PubMed, was conducted from 2003 through December 2011, utilizing the following search terms: (((“Lumbosacral Region”[MeSH] OR “Lumbar Vertebrae”[MeSH]) AND “Spinal Fusion”[MeSH]) OR “lumbar fusion”[All Fields] OR (“lumbar”[title] AND “fusion”[title])) AND (interbody) AND (low back pain). The search yielded 183 citations. Clinical series reported in English-language journals dealing with adult patients who had undergone fusion with instrumentation for degenerative lumbar disease were selected. Relevant articles pertaining to the comparison of interbody fusion techniques with other surgical techniques or nonsurgically treated controls were selected and are summarized in Table 1. A number of case series provide supporting data and are referenced in the bibliography.

Scientific Foundation

Recent trends in spinal surgery involve the use of interbody fusion techniques, including ALIF, posterior lumbar interbody fusion (PLIF), transforaminal lumbar interbody fusion (TLIF), or axial lumbar interbody fusion as a means to enhance the rate of successful arthrodesis. Authors of several studies have compared the results of these techniques with respect to each other as well as with respect to PLF.

Comparison of Interbody Fusion and PLF

Christensen et al. reported a series of 148 patients with severe low-back pain who were prospectively randomized to treatment with PLF with pedicle screws or ALIF with Brantigan cages in addition to posterior instrumentation and PLF.[2] The Dallas Pain Questionnaire and the Low Back Pain Rating Scale were used to assess outcomes. Patients treated with circumferential procedures had better overall functional outcome, but this was not statistically significant ($p = 0.08$). This patient group did have statistically significant less leg pain at the 1-year follow-up evaluation ($p < 0.03$) and less maximum back pain at 2 years ($p < 0.04$). Fusion rate, which was determined on static plain radiographs, was significantly higher in the circumferential fusion group (92%) than in the PLF with pedicle screws group (80%) ($p < 0.04$). The circumferential fusion group had an 82% interbody fusion rate. The reoperation rate was significantly lower in the circumferential group (7%) than in the PLF group (22%) ($p < 0.009$). This paper provides Level II evidence supporting the role of interbody grafts in improving arthrodesis rates and the role of interbody grafts in improving outcome with respect to back and leg pain. The lack of flexion-extension views or CT scans to supplement the static radiographs rendered a less accurate evaluation of fusion status, and thus this study was downgraded to Level II.

Kim et al. also performed a prospective randomized study comparing PLF, PLIF, and PLIF+PLF in 167 patients who underwent 1- or 2-level fusion surgery for degenerative lumbar disease.[7] The patients were randomized into one of 3 treatment groups: Group 1 (PLF; $n = 62$), Group 2 (PLIF; $n = 57$), and Group 3 (PLF+PLIF; $n = 48$). The minimum follow-up was 3 years. Local autograft from the lamina and spinous processes was placed in the interbody cage, and iliac crest autograft was used for PLF. Clinical follow-up included the visual analog scale (VAS), Oswestry Disability Index (ODI), and Kirkaldy-Willis criteria. Radiological follow-up included flexion-extension radiographs and a CT scan when fusion status was in question. All groups demonstrated significant clinical improvement from preoperative status. There was no significant difference in clinical results or fusion rates (92% in Group 1, 95% in Group 2, and 96% in Group 3; $p > 0.05$) between the 3 groups. The PLIF group had better sagittal balance than the instrumented PLF. With the addition of PLF to the PLIF, the patients reported donor site pain as well as increased blood loss and operative time, all of which were secondary to harvesting iliac crest. The authors suggested that the addition of PLF is not beneficial when PLIF is performed. This study provides Level II evidence against the addition of PLF to PLIF. The study was downgraded to Level II because of a lack of power analysis and no report of the rate of loss to follow-up.

Greenough et al. reported a prospective case series assessing the results of instrumented PLF in 135 patients

TABLE 1: Interbody techniques for lumbar fusion: summary of evidence*

Authors & Year	Level of Evidence	Description	Results	Conclusion
Christensen et al., 2002	II: Prospective randomized clinical trial that was downgraded due to using only static radiographs to evaluate fusion status.	A prospective randomized clinical study analyzed the effects of circumferential fusion using ALIF radiolucent carbon fiber cages & titanium posterior instrumentation vs instrumented PLF (w/ pedicle screws) w/ 2-yr follow-up.	The circumferential lumbar fusion group had a higher fusion rate w/ significantly fewer reops, showed a tendency toward better functional outcome than the instrumented PLF group.	The authors favored circumferential fusion as a definitive surgical procedure in complex lumbar pathology involving major instability, flat back, & previous disc surgery in younger pts, compared w/ PLF w/ pedicle screws alone.
Fritzell et al., 2002	II: Prospective comparison study downgraded to Level III due to using only static radiographs to evaluate the fusion status.	A multicenter randomized study to compare 3 commonly used surgical techniques to achieve lumbar fusion in pts w/ severe chronic low-back pain due to disc degeneration or spondylosis.	All surgical techniques (PLF, PLF combined w/ pedicle screw fixation, & PLF combined w/ pedicle screw placement & interbody fusion using ALIF or PLIF) were found to reduce pain & decrease disability substantially.	All fusion techniques used in the study could reduce pain & improve function in pts w/ severe chronic low-back pain. Fusion rate was significantly better when internal fixation was used.
Kim et al., 2006	II: Study downgraded to Level II because of a lack of power analysis & no report of loss to follow-up.	A prospective randomized study compared 3 fusion methods: instrumented PLF, PLIF, & PLF+PLIF w/ minimum 3-yr follow-up.	No statistical differences were found among the 3 groups in terms of clinical outcomes & pseudarthrosis rates.	No significant differences in clinical results & union rates were found among the 3 fusion methods. PLIF had better sagittal balance than PLF. PLIF w/o PLF had advantages of the elimination of donor site pain, shorter operating time, & less blood loss.
Thalgott et al., 2009	II: This is a randomized control study downgraded to Level II due to lack of power analysis.	A prospective, randomized clinical trial compared the outcomes & fusion rates of an ALIF procedure w/ freeze-dried or frozen FRA w/ a minimum 24-mo follow-up.	The freeze-dried graft had a higher likelihood of pseudarthrosis.	When the results are considered in terms of clinical outcomes, the 2 methods of graft preservation perform w/ few statistically significant differences. Radiographic analysis showed that the freeze-dried graft had a higher likelihood of pseudarthrosis.
Videbaek et al., 2011	II: Downgraded because this study is not an actual randomized controlled trial.	A randomized clinical trial compared ALIF+instrumented PLF vs instrumented PLF in pts w/ severe back pain w/ 10-yr follow-up.	Sagittal balance parameters were similar btwn randomization groups. None of the parameters differed significantly btwn pts w/ an ODI from 0 to 40 & pts w/ ODI >40. Balanced pts had a significantly superior outcome as measured by ODI than unbalanced pts.	No difference in the investigated sagittal balance parameters btwn pts treated w/ PLF+ALIF vs those w/ instrumented PLF alone.
Abdu et al., 2009	III: A retrospective cohort comparison study w/ a lack of fusion status evaluation & possible selection bias as the surgeon chose the fusion technique.	This study retrospectively examined data collected during a prospective, randomized trial of 380 pts w/ degenerative spondylolisthesis & stenosis treated w/ standard decompressive laminectomy & 1 of 3 fusion techniques at the surgeon's discretion: PLF, PPS, or PPS+interbody fusion.	Early outcomes varied, favoring PLF compared w/ PPS & PPS compared w/ 360° at 6 wks & 3 mos. At 2 yrs, 360° had better outcomes. However, these differences were not maintained at 3- & 4-yr follow-up, when there were no statistically significant differences btwn the 3 fusion groups.	In pts w/ degenerative spondylolisthesis & associated spinal stenosis, no consistent differences in clinical outcomes were seen among fusion groups over 4 yrs.

(continued)

TABLE 1: Interbody techniques for lumbar fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Results	Conclusion
Pradhan et al., 2002	III	Retrospective review of pts who underwent either an anterior interbody or posterolateral intertransverse process w/ single-level instrumented lumbar spinal fusion performed.	There was significantly less blood loss, need for transfusion, amount of blood transfused, operative time, & hospital stay for pts w/ anterior fusion procedures.	The anterior approach to single-level lumbar fusion is associated w/ less morbidity than the posterolateral approach. However, both approaches to single-level lumbar fusion produce similar early fusion rates & clinical results.
Schofferman et al., 2001	III	A prospective randomized comparison of ALIF+transpedicular instrumentation+PLF (360° fusion) to ALIF+transpedicular instrumentation w/o PLF (270° fusion) w/ an average follow-up of 35 mos.	There were significant postop improvements in pain & function in both groups w/o significant differences in percentage solid ALIF. However, the 270° fusion group had significantly less blood loss, shorter operative times, shorter LOS, & lower professional fees.	Both the 360° & 270° fusions significantly reduce pain & improve function, & there are no significant clinical differences btwn them. There were shorter operating times, less blood loss, lower costs, & less utilization of health care resources associated w/ the 270° fusions.
Yan et al., 2008	III: Retrospective comparison study	This study retrospectively compared PLIF & TLIF w/ pedicle screw fixation in pts w/ degenerative spondylolisthesis.	All pts had bone fusion, & there were no cases of cage extrusion. The JOA score in all pts was good or excellent. Both techniques achieve statistical significance in restoration of disc & foraminal; however, there was no statistical difference btwn the 2 techniques.	Interbody fusion w/ either a PLIF technique or a TLIF technique provides good outcomes in the treatment of adult degenerative spondylolisthesis.
Greenough et al., 1998	III	A prospective case series assessed the results of instrumented PLF in pts w/ intractable back pain w/ a minimum 2-yr follow-up & compared them w/ a historical control of ALIF.	The method of outcome assessment profoundly affected the results; whereas 65% of pts rated themselves significantly improved by the procedure, only 19% achieved a good or excellent result on the LBOS.	Overall, the results of instrumented PLF were inferior to those in a similar series treated by ALIF.

* ALIF = anterior lumbar interbody fusion; FRA = femoral ring allograft; JOA = Japanese Orthopaedic Association; LBOS = Low Back Outcome Score; LOS = length of stay; ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; PLIF = posterior lumbar interbody fusion; PPS = instrumented PLF with pedicle screws; pts = patients; TLIF = transforaminal lumbar interbody fusion.

with intractable back pain who were treated by a single surgeon.[6] They compared the results of this cohort with a previously published historical control of 151 patients who underwent ALIF performed by the same single surgeon. A solid bony fusion was obtained in 82% of patients as assessed mainly using static radiographs. The Low Back Outcome Score was statistically significantly better in the historical cohort of ALIF patients than in the instrumented PLF group ($p < 0.01$). This report provides Level III evidence that ALIF has better clinical outcomes than instrumented PLF in patients with chronic back pain. However, the authors did not compare the fusion rates between the 2 fusion techniques, and they used a historical ALIF cohort to compare the clinical results.

Videbaek et al. studied patient cohorts from a prospective randomized study analyzing the long-term (8–13 years) impact of ALIF+PLF versus PLF on sagittal spinal balance in 1- or 2-level fusion surgery.[11] The original study patients underwent additional radiography, which is the focus of this paper. There were 48 patients in the ALIF+PLF group and 44 in the PLF group. Posterolateral fusion was performed with pedicle screw fixation and iliac crest bone graft in the PLF group and with pedicle screw fixation or facet screw fixation in the ALIF+PLF group, depending on the necessity of posterior decompression. In the ALIF+PLF group, the PLF was performed first followed by ALIF in one stage. The radiographic parameters included pelvic incidence, sacral slope, pelvic tilt, maximal thoracic kyphosis, maximal lumbar lordosis, and segmental lordosis. The clinical outcome assessed was ODI. All parameters except for segmental lordosis showed no statistical difference in the 2 groups. Patients with 2-level fusion were over-represented in the ALIF+PLF group. The difference in segmental lordosis was eliminated in subgroup analysis according to number of levels fused. There was a significant positive correlation between lumbar lordosis and ODI score ($r = 0.31$, $p < 0.01$) when considering the entire cohort. The authors concluded that the sagittal alignment is not dependent on anterior column support and lumbar lordosis correlated with postoperative outcome. This paper did not focus on fusion status and instead focused on the sagittal balance and radiographic alignment parameters. The authors asked participants of a prior prospective, randomized trial to undergo new imaging studies. The follow-up rate was less than 65%, and therefore the report was downgraded to Level II evidence. This paper is a subsequent analysis of a prospective, randomized trial.

Schofferman et al. reported a prospective, randomized study comparing 26 patients who were treated with ALIF+pedicle screws+PLF (360° fusion group) with 22 patients who were treated with ALIF+pedicle screws without PLF (270° fusion group).[9] An FRA filled with cancellous allograft chips is used in ALIF. Flexion-extension plain radiographs were used to evaluate fusion status. The mean follow-up period was 35 months. Clinical outcomes were measured using the Numerical Rating Scale and the ODI. In the 360° fusion group, the PLF part of the procedure failed to heal 68% of the time. There was no significant difference ($p = 0.6$, chi-square test) in the fusion rate of the interbody graft between the groups, although there was a trend favoring the 270° fusion group (77% fusion rate in the 360° fusion group compared with 89% fusion rate in the 270° fusion group). The 270° fusion group had a shorter operating time, less intraoperative blood loss, and shorter length of stay (LOS) (all $p < 0.05$). This study provides Level III evidence that the addition of PLF to an ALIF with pedicle screw construct increases blood loss, LOS, and operating time without any resultant benefit. It was downgraded due to a lack of power analysis and suboptimal randomization. In addition, the patient population was not well defined.

Abdu et al. reported a subgroup analysis of 3 different fusion methods from data collected during a prospective randomized trial of 395 surgically treated patients with degenerative spondylolisthesis and stenosis.[1] In addition to decompressive laminectomy, one of 3 fusion techniques was used at the surgeon's discretion: in situ PLF; instrumented PLF with pedicle screws (PPS); or PPS plus interbody fusion using ALIF, TLIF, or PLIF (360° fusion). Main outcome measures were the 36-Item Short Form Health Survey (SF-36) bodily pain and physical function scales and the modified ODI assessed at 6 weeks, 3 months, 6 months, and yearly to 4 years. From the surgical cohort, 380 patients (96%) met inclusion criteria for analysis. The distribution of surgical procedures was as follows: 21% ($n = 80$) underwent PLF; 56% ($n = 213$) underwent PPS; 17% ($n = 63$) underwent 360° fusion; and 6% ($n = 23$) underwent a decompression without a fusion. Significant differences in outcome were observed that varied during the early follow-up period. Greater improvements in the physical function score were observed for PLF compared with PPS at 6 weeks (physical function: 12.73 vs 6.22, $p < 0.020$) and 3 months (physical function: 25.24 vs 18.95, $p < 0.025$). More substantial improvements in the ODI scores were observed for patients undergoing PPS compared with the 360° fusion cohort at 6 weeks (ODI: −14.46 vs −9.30, $p < 0.03$) and 3 months (ODI: −22.30 vs −16.78, $p < 0.02$). At 2 years, the 360° fusion cohort demonstrated statistically significant improvement in bodily pain and physical function scores compared with the PLF cohort ([bodily pain: 39.08 vs 29.17, $p < 0.011$] and [physical function: 31.93 vs 23.27, $p < 0.021$]) and the PPS cohort ([bodily pain: 39.08 vs 29.13, $p < 0.002$] and [physical function: 31.93 vs 25.29, $p < 0.036$]). The differences in outcome between the 3 fusion cohorts were not observed beyond 2 years, with no significant differences at either the 3- or 4-year follow-up time point. The authors concluded that there was no significant advantage of one fusion technique over another on clinical outcomes at 4-year follow-up; however, longer follow-up may be needed. This report is a subgroup analysis of varied fusion methods using the combined cohorts from a randomized controlled trial and a concurrent observational cohort. It is not an actual randomized controlled trial itself but rather a prospective comparison study (Level II) with a lack of fusion status evaluation. Another limitation of this report is that the fusion techniques were not randomly assigned and thus selection bias may exist since the surgeons chose which technique to use at their own discretion. Thus, it is downgraded to Level III.

Fritzell et al. performed a randomized, prospective, multicenter trial involving 294 patients with chronic low-back pain due to degenerative disc disease at 1 or 2 levels.[4,5] Patients were randomized to one of 4 treatment groups. Patients in Group 1 (73 patients) underwent a noninstrumented PLF. Those in Group 2 (74 patients) were treated with PLF with pedicle screw fixation; patients in Group 3 were treated with interbody arthrodesis supplemented with pedicle screw fixation (56 of these patients underwent ALIF with pedicle screws and 19 of these patients underwent PLIF with PLF and pedicle screws). Group 4 was treated nonsurgically. Ninety-one percent of patients were available for follow-up by an independent observer. Although all surgical groups did substantially better than the nonsurgical group, there were no statistically significant differences in ODI, Low Back Pain Questionnaire, Million VAS, and General Function Score between the surgical groups. The early complication rate was 6% in Group 1, 16% in Group 2, and 31% in Group 3. The fusion rate was evaluated on plain radiographs (without flexion-extension views) and was 72% in Group 1, 87% in Group 2, and 91% in Group 3. The authors concluded that all surgical groups had similar functional outcomes, but they noted that their study did lack power to detect a difference in functional outcome between the surgical groups. There was an increase in the fusion rate in the instrumented group and in the interbody group compared with the noninstrumented group ($p = 0.004$). This paper provides Level II evidence supporting the beneficial effects of instrumentation and interbody grafts on fusion rates. However, the fusion status is determined by static radiographs. It is downgraded due to lack of power to detect a difference in functional outcomes and also due to the use of only static radiographs to evaluate fusion status.

With respect to complication rates, the same authors found that overall complication rates were higher in the instrumented PLF and interbody groups than in the noninstrumented PLF group.[3] The early complication rate was 6% in the PLF group, 18% in the PLF with screw group, and 31% in the 360° fusion group ($p = 0.001$). There was no significant difference in the reoperation rate between the interbody group and the PLF with pedicle screw group. These reoperations would appear to be unrelated to the use of an interbody implant. Seventeen of the 29 complications reported in the 360° fusion group did not necessarily result from the interbody procedure itself. These complications included donor site pain, pressure sores, and screw malposition. Four complications were specifically related to the anterior approach: 2 iliac vein lacerations and 2 sympathetic nerve injuries. There were 7 instances of new nerve root pain, 2 of which required reoperation within 2 years. The 2-year follow-up complication rate was 12% in the PLF group, 22% in the PLF with screws group, and 40% in the 360° fusion group ($p = 0.0003$). This complication rate includes reoperations for instrumentation removal, whether the removal was performed because of any problems associated with the instrumentation. The only delayed complication reported in the interbody group was continued donor site pain in the patients who underwent ALIF. The lack of beneficial effect on functional outcome, along with the higher complication rate associated with the circumferential procedures, may be interpreted as evidence against the use of circumferential procedures as a means to improve patient outcomes.

Pradhan et al. performed a retrospective review to compare 58 patients who were treated with lumbar ALIF with BAK cages (Sulzer Spine-Tech) (Group 1) with 64 patients who were treated with PLF with pedicle screw fixation (Group 2).[8] The follow-up period was 22 months for ALIF and 26 months for PLF. Fusion was assessed based on flexion-extension radiographs and CT scanning for ambiguous cases. Radiographic fusion was confirmed in 95% of the Group I patients and in 92% of the Group II patients; however, this difference was not statistically significant. The ALIF cohort had a lower operative blood loss, shorter operative time, and shorter LOS ($p < 0.01$). The complication rates or clinical outcomes were not statistically different between the groups. Although this paper provides Level III evidence indicating that placement of an interbody graft through a stand-alone ALIF technique does not improve fusion rates compared with PLF, the small size of the treatment groups in this study makes any statement regarding functional outcomes suspect. The ALIF group was reported to have a shorter LOS, less blood loss, and less exposure to anesthetic agents.

Implants Used for Interbody Fusion

Thalgott et al. performed a prospective, blinded, randomized, single-site study from a single surgeon's patient population to evaluate the clinical and radiographic outcome differences between frozen and freeze-dried FRA for ALIF as part of a circumferential fusion for the treatment of degenerative disc disease including Grade I degenerative spondylolisthesis.[10] Patients were observed for a minimum of 24 months. Outcome measures included complications, fusion status, implant intactness, 1–10 pain scores, ODI, and SF-36 scores. Radiographic assessment was performed by an independent, blinded, board-certified radiologist and included dynamic lateral radiographs as part of the fusion assessment. The ODI improved more than 10 points in 62.5% of patients and SF-36 scores improved more than 10 points in 27.5% of patients. There was no statistically significant difference in clinical outcomes between the 2 groups. However, the freeze-dried allograft had a statistically higher rate of pseudarthrosis ($p = 0.026$). This paper suggests that frozen FRA has a lower rate of pseudarthrosis compared with freeze-dried allograft. In this study, the patients with 100% of their treated levels fused had better clinical outcomes than patients with pseudarthrosis. These differences were statistically significant with regard to the SF-36 Physical Component Summary and trended toward significance with the ODI. This study did not have a power analysis; it was therefore downgraded to Level II evidence in support of the use of frozen FRA instead of freeze-dried allograft for use in anterior lumbar fusion procedures.

Yan et al. performed a retrospective review of 187 patients who underwent either a PLIF with bilateral cages or a TLIF with unilateral placement of an interbody cage for the treatment of single-level degenerative spondylolis-

thesis.[12] Ninety-one patients underwent PLIF with 2 cages and pedicle fixation (Group 1), and 96 patients underwent TLIF with 1 cage and pedicle fixation (Group 2). Before surgery and at the 2-year follow-up, pain and functional disability were quantified using the VAS and Japanese Orthopaedic Association scales, respectively. The follow-up rate was 93.4% (85 of 91 patients) in the PLIF group and 94.8% (91 of 96 patients) in the TLIF group. All patients had bone fusion, and there were no cases of cage extrusion. Both groups demonstrated similar clinical and radiographic outcomes. The authors concluded that interbody fusion with either a PLIF technique or a TLIF technique provides good outcomes in the treatment of adult degenerative spondylolisthesis. The TLIF procedure is simpler and is as safe and effective as the PLIF technique. This study provides Level III evidence supporting TLIF over PLIF as a lumbar fusion option.

Summary

The medical evidence continues to suggest that interbody techniques are associated with higher fusion rates compared with PLF in patients with degenerative spondylolisthesis who demonstrate preoperative instability. However, there is no conclusive evidence supporting better clinical and radiographic outcomes based on different interbody fusion techniques. The evidence generally comprises Level II and III studies.

The addition of PLF when PLIF or ALIF is performed is optional and has been found to be associated with increased cost and complications.

With regard to type of interbody spacer used, frozen ALIF allograft is associated with lower pseudarthrosis rates compared with freeze-dried ALIF allograft. This is a Grade C recommendation supported by a single Level II study.

There is no conclusive evidence supporting better clinical or radiographic outcomes based on technique when performing interbody fusion. No general recommendation can therefore be made regarding the technique that should be used to achieve interbody fusion. We did not analyze any comparisons of minimally invasive surgery versus traditional open surgery in this report.

Key Issues for Future Investigation

1) The optimal approach and technique for interbody fusion at different levels of the lumbar spine should be investigated using prospective comparison/cohort studies to ascertain which one has the lowest complication rate along with the highest fusion rate and greatest clinical outcomes benefit.

2) The cost-effectiveness and long-term outcomes of different techniques for lumbar fusion should be investigated.

A prospectively registered database will assist in reporting the efficacy and associated complications of new approaches.

Acknowledgments

We would like to acknowledge the AANS/CNS Joint Guidelines Committee (JGC) and Tsung His Tu, M.D., 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., 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 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: Mummaneni. 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: Mummaneni. Study supervision: Kaiser.

References

1. Abdu WA, Lurie JD, Spratt KF, Tosteson AN, Zhao W, Tosteson TD, et al: Degenerative spondylolisthesis: does fusion method influence outcome? Four-year results of the spine patient outcomes research trial. **Spine (Phila Pa 1976) 34:**2351–2360, 2009
2. Christensen FB, Hansen ES, Eiskjaer SP, Høy K, Helmig P, Neumann P, et al: Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. **Spine (Phila Pa 1976) 27:**2674–2683, 2002
3. Fritzell P, Hägg O, Nordwall A: Complications in lumbar fusion surgery for chronic low back pain: comparison of three surgical techniques used in a prospective randomized study. A report from the Swedish Lumbar Spine Study Group. **Eur Spine J 12:**178–189, 2003
4. Fritzell P, Hägg O, Wessberg P, Nordwall A: 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. **Spine (Phila Pa 1976) 26:**2521–2534, 2001
5. Fritzell P, Hägg O, Wessberg P, Nordwall A: Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish Lumbar Spine Study Group. **Spine (Phila Pa 1976) 27:**1131–1141, 2002

6. Greenough CG, Peterson MD, Hadlow S, Fraser RD: Instrumented posterolateral lumbar fusion. Results and comparison with anterior interbody fusion. **Spine (Phila Pa 1976) 23:**479–486, 1998
7. Kim KT, Lee SH, Lee YH, Bae SC, Suk KS: Clinical outcomes of 3 fusion methods through the posterior approach in the lumbar spine. **Spine (Phila Pa 1976) 31:**1351–1358, 2006
8. Pradhan BB, Nassar JA, Delamarter RB, Wang JC: Single-level lumbar spine fusion: a comparison of anterior and posterior approaches. **J Spinal Disord Tech 15:**355–361, 2002
9. Schofferman J, Slosar P, Reynolds J, Goldthwaite N, Koestler M: A prospective randomized comparison of 270 degrees fusions to 360 degrees fusions (circumferential fusions). **Spine (Phila Pa 1976) 26:**E207–E212, 2001
10. 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
11. Videbaek TS, Bünger CE, Henriksen M, Neils E, Christensen FB: Sagittal spinal balance after lumbar spinal fusion: the impact of anterior column support results from a randomized clinical trial with an eight- to thirteen-year radiographic follow-up. **Spine (Phila Pa 1976) 36:**183–191, 2011
12. Yan DL, Pei FX, Li J, Soo CL: Comparative study of PILF and TLIF treatment in adult degenerative spondylolisthesis. **Eur Spine J 17:**1311–1316, 2008

Manuscript submitted March 13, 2014.
Accepted April 8, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14276.
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.

J Neurosurg Spine 21:75–78, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 12: Pedicle screw fixation as an adjunct to posterolateral fusion

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

[1]Department of Neurosurgery, Brigham and Women's Hospital, Boston, Massachusetts; [2]Department of Neurosurgery, University of Utah, Salt Lake City, Utah; [3]Alan and Jacqueline Stuart Spine Research Center, Department of Neurosurgery, Lahey Clinic, Burlington, and Tufts University School of Medicine, Boston, Massachusetts; [4]Department of Neurosurgery, University of Wisconsin, Madison, Wisconsin; [5]Bone and Joint Clinic of Houston, Houston, Texas; [6]Department of Neurological Surgery, University of California, San Francisco, California; [7]Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, New York, New York; [8]Center for Sports Medicine and Orthopaedics, Chattanooga, Tennessee; [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 utilization of pedicle screw fixation as an adjunct to posterolateral lumbar fusion (PLF) has become routine, but demonstration of a definitive benefit remains problematic. The medical evidence indicates that the addition of pedicle screw fixation to PLF increases fusion rates when assessed with dynamic radiographs. More recent evidence, since publication of the 2005 Lumbar Fusion Guidelines, suggests a stronger association between radiographic fusion and clinical outcome, although, even now, no clear correlation has been demonstrated. Although several reports suggest that clinical outcomes are improved with the addition of pedicle screw fixation, there are conflicting findings from similarly classified evidence. Furthermore, the largest contemporary, randomized, controlled study on this topic failed to demonstrate a significant clinical benefit with the use of pedicle screw fixation in patients undergoing PLF for chronic low-back pain. This absence of proof should not, however, be interpreted as proof of absence. Several limitations continue to compromise these investigations. For example, in the majority of studies the sample size is insufficient to detect small increments in clinical outcome that may be observed with pedicle screw fixation. Therefore, no definitive statement regarding the efficacy of pedicle screw fixation as a means to improve functional outcomes in patients undergoing PLF for chronic low-back pain can be made. There appears to be consistent evidence suggesting that pedicle screw fixation increases the costs and complication rate of PLF. High-risk patients, including (but not limited to) patients who smoke, patients who are undergoing revision surgery, or patients who suffer from medical conditions that may compromise fusion potential, may appreciate a greater benefit with supplemental pedicle screw fixation. It is recommended, therefore, that the use of pedicle screw fixation as a supplement to PLF be reserved for those patients in whom there is an increased risk of nonunion when treated with only PLF.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14277)

Key Words • **lumbar spine** • **pedicle screw** • **posterolateral fusion** • **adjunct** • **practice guidelines**

Recommendations

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

Grade B

Pedicle screw fixation is recommended when posterolateral lumbar fusion (PLF) is used to manage low-back pain in patients at high risk for pseudarthrosis.

Routine use of pedicle screw fixation as an adjunct to PLF for patients with degenerative disc disease is an option. There is consistent evidence that the use of pedicle screws enhances the fusion rate; however, a positive correlation with respect to clinical outcome has not been consistently demonstrated.

Abbreviations used in this paper: DPQ = Dallas Pain Questionnaire; JOA = Japanese Orthopaedic Association; PLF = posterolateral lumbar fusion; SF-36 = 36-Item Short Form Health Survey.

Rationale

Arthrodesis of the lumbar spine has become an accepted treatment option for spinal disorders manifesting with low-back pain. Although there is an ever-increasing collection of techniques to achieve a successful arthrodesis, the traditional PLF remains a commonly performed and successful surgical approach. The inclusion of internal fixation through pedicle screw stabilization has become a routine addition to PLF. Pedicle screw fixation as an adjunct to PLF is known to have advantages, including a higher fusion rate, and disadvantages, including higher cost and a higher rate of complications. The purpose of this update is to review the current medical literature and determine if the evidence supports or refutes the role for pedicle screws as an adjunct of PLF in the treatment of degenerative spinal disorders, such as low-grade degenerative spondylolisthesis, leading to low-back pain.

Search Criteria

A computerized search of the National Library of Medicine database of the literature published from July 2003 to December 2011 was performed using the following search terms: ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ("low back pain"[MeSH] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields]) AND ("Bone Screws"[MeSH] OR "pedicle screw*"[All Fields] AND (("2003"[PDAT]: "3000"[PDAT]) AND "humans"[MeSH] AND English[lang])) AND ("humans"[MeSH] AND English[lang] AND ("aged"[MeSH] OR "aged, 80 and over"[MeSH])). The search was limited to clinical series reported in English-language journals dealing with adult patients who had fusion with instrumentation for degenerative lumbar disease and yielded 258 publications. Among the articles reviewed, references were included if they described a comparison of fusion techniques with or without instrumentation. These references are summarized in Table 1.

Scientific Foundation

There is a wealth of literature demonstrating the positive impact of pedicle screw fixation on fusion rates in patients treated with PLF. Although a small number of papers report an improvement in functional outcomes with pedicle screw fixation, the quality of these data is low from an evidence-based medicine perspective.[9,13] The results of the articles reviewed indicates that pedicle screw fixation for degenerative spondylosis has little if any impact on functional outcome.[5,6,9,11] This conclusion served as the basis for the recommendations of the previous Lumbar Fusion Guidelines.[10] Since our original review there have been several well-designed studies that address the utility of pedicle screw fixation in the context of degenerative disc disease of the lumbar spine.

Korsgaard et al. performed a randomized prospective study evaluating the impact of pedicle screws with respect to clinical outcome in 130 patients undergoing treatment of degenerative lumbar disease.[8] All patients underwent PLF and were randomly assigned to either a noninstrumented or instrumented cohort. Fusion status was assessed using the Christensen classification, which utilizes static anteroposterior and lateral radiographs.[3] Clinical outcomes were evaluated using the Dallas Pain Questionnaire (DPQ). There were no significant differences between the treatment cohorts with respect to baseline demographic characteristics. At 2 years after surgery, no significant difference was observed between the 2 groups with respect to fusion rate or clinical outcome. Bjarke Christensen et al. reevaluated this same group of patients 5 years after surgery and found no significant difference in functional outcome; however, the authors did observe a higher reoperation rate in the instrumented group (25% vs 14% in the noninstrumented group).[2] It should be recognized, however, that only 11% of the reoperations in the instrumented group were for complications associated with the hardware. A subgroup analysis demonstrated that patients with "primary degenerative instability" experienced a greater improvement on the DPQ with instrumentation as compared with the noninstrumented cohort.

Andersen et al. performed a prospective nonrandomized study evaluating the role of pedicle screw fixation in patients over 60 years of age undergoing a posterolateral fusion with fresh-frozen allograft for degenerative lumbar spondylosis.[1] Pedicle screw stabilization was performed at the discretion of the operating surgeon. The authors used allograft in an attempt to avoid the morbidity associated with harvesting iliac crest autograft. The indications for a fusion included preoperative or anticipated iatrogenic instability, as well as significant back pain before surgery. Clinical outcome was assessed with the DPQ. Fusion status was assessed with static plain radiographs. All outcome measures were improved with instrumentation compared with noninstrumented fusion. The fusion rate was higher in the instrumented group (81% vs 68%). It should be remembered that the study was not randomized and the mean age of the patients in the instrumented group was lower than the mean age of the patients in the noninstrumented group.

Several case series have also provided evidence regarding PLF for degenerative lumbar spondylosis.[7] Epstein investigated the outcome in 75 cases involving geriatric patients who underwent noninstrumented lumbar fusion with local autograft and a beta-tricalcium phosphate graft extender.[4] Clinical outcome was assessed with the 36-Item Short Form Health Survey (SF-36), and fusion was assessed with CT scans and flexion-extension radiographs. In this study, Epstein documented a fusion rate of 83% and an improvement in all aspects of the SF-36, with the exception of mental health, which remained unchanged.

Tsutsumimoto et al. performed a retrospective analysis of a series of 42 cases involving patients who underwent noninstrumented PLF for degenerative lumbar stenosis.[12] Fusion status was assessed with flexion-extension radiographs, and clinical outcome was measured with the Japanese Orthopaedic Association (JOA) scale. The fusion rate was 74%. At 5 years postsurgery there was a significant improvement in the JOA scores of the patients in whom fusion was achieved when compared with those who had

TABLE 1: Pedicle screw fixation as an adjunct to PLF: summary of evidence*

Authors & Year	Level of Evidence	Brief Description	Comments
Korsgaard et al., 2002	II	Prospective randomized study of 130 pts w/ degenerative lumbar spondylosis. Pts underwent PLF w/ or w/o PS fixation. Follow-up 2 yrs. DPQ used for outcome assessment. Lumbar lordosis & fusion determined by plain radiographs. There was no significant btwn-groups difference on DPQ. No correlation btwn lordosis & DPQ. Fusion rate similar w/ or w/o PS fixation.	No power calculation. Static radiographs used for fusion analysis. Nonstandard, divergent method of sacral screw insertion.
Andersen et al., 2009	III	Prospective cohort study of 94 pts older than 60 yrs of age who underwent PLF w/ allograft. No instrumentation was used in 51 cases; PS fixation was used in 43. Outcome was assessed using DPQ, LBPRS, & SF-36. Fusion was assessed using plain radiographs. Pts were followed for 2–7 yrs. Pts treated w/ PS fixation had superior outcome (mean follow-up 4.3 yrs).	Downgraded to Level III because fusion was assessed w/ static radiographs & the follow-up rate was 76%.
Jäger et al., 2003	III	Prospective cohort study of 33 pts. All underwent PLF; instrumentation was used in 17 cases. Indication for surgery defined only as degenerative instability. Fusion was assessed w/ standard radiographs. Flexion-extension or CT was used only if needed. ODI was used. No difference reported in fusion or clinical outcomes. Pt accrual required 11 yrs, creating potential for substantial bias.	Limitations included small sample size & lack of validated standard for evaluating radiographic evidence of fusion. Downgraded to Level III evidence.
Bjarke Christensen et al., 2002	II	Prospective randomized study of 129 pts w/ chronic low-back pain. Pts were treated w/ PLF w/ or w/o PS fixation & followed for 5 yrs (93% follow-up). DPQ & LBPRS were used. For the entire cohort there were no statistically significant differences in functional outcome or fusion rates. Fusion was assessed w/ static radiographs. Subgroup analysis demonstrated that pts w/ isthmic spondylolisthesis had improved outcomes w/ noninstrumented PLF while pts w/ primary degenerative instability had better outcomes w/ instrumented PLF.	Block randomization, w/ power analysis. Static radiographs used for fusion analysis. Unclear if a standardized surgical technique utilized. LBPRS was not administered prior to surgery.
Fischgrund et al., 1997	II	Prospective randomized study of 76 pts w/ spondylolisthesis & spinal stenosis. Pts were randomized to PLF w/ or w/o PS fixation. Fusion rate was higher in instrumented group (82% vs 45%), while outcome was superior in noninstrumented group (85% vs 76%).	Small sample size & nonvalidated outcome & fusion measures. Follow-up 88% at 2 yrs.
Fritzell et al., 2002	II	Prospective randomized study of 222 pts randomized to PLF, PLF + PS fixation, & PLF + PS + IBF. Follow-up 91% at 2 yrs. All groups improved equally on VAS & ODI. Complication rates were 6%, 16%, & 31%.	No power calculation. Underpowered.
Lorenz et al., 1991	II	Prospective randomized study of 68 pts w/ disabling back pain. Pts were randomized to PLF or PLF + PS fixation. Follow-up at mean 26 mos w/ flexion-extension radiographs & RTW. Fusion rate, pain score, & RTW superior w/ PS fixation.	RTW & pain score. Lack of validated outcome measure.
Zdeblick, 1993	II	Prospective, randomized study of 124 pts: PLF, PLF + semi-rigid PS fixation, PLF + PS. Fusion determined w/ flexion-extension radiographs at 1 yr: 65%, 77%, 95%.	Clinical outcome measure not validated.

* DPQ = Dallas Pain Questionnaire; IBF = interbody fusion; LBPRS = Low Back Pain Rating Scale; ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; PS = pedicle screw; pt = patient; RTW = return to work; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale.

pseudarthrosis (3.5 vs 2.5). Regression analysis revealed that fusion status and comorbidity were the strongest predictors of the improvement demonstrated on the JOA scale.

Summary

The role of pedicle screw stabilization as an adjunct to PLF for lumbar degenerative disease continues to be an area of intense investigation. In the years since the original guideline publication, new evidence has been generated, demonstrating that the improved fusion rate with the use of pedicle screws can lead to improved clinical outcomes (Level II) and that pseudarthrosis is associated with worse long-term clinical outcome (Level IV). An improved fusion rate with the application of pedicle screw stabilization has been well established from previous published reports. Although the recent literature is more suggestive of a relationship between successful fusion and improved clinical outcomes, a direct clinical benefit for the use of pedicle screws still has not been conclusively established. We therefore recommend that pedicle screws be used routinely as an adjunct to PLF for low-back pain only in cases that pose an increased risk for pseudarthrosis. Those cases include, but are not limited to, those involving patients who smoke, present with kyphotic deformity, or suffer systemic diseases associated with poor bone healing. The use of pedicle screw fixation in other cases is associated with an increase in the fusion rate, but any association with improved outcome is less well defined.

Key Issues for Future Investigation

There is convincing support in the literature for the beneficial impact of pedicle screw fixation on arthrodesis. There is also support for the beneficial impact of a successful arthrodesis on clinical outcome. Nonetheless, studies examining the impact of pedicle screw fixation on clinical outcome have been inconclusive. Further investigation should elucidate the cause of this apparent contradiction. Possible explanations include the complication profile of pedicle screw insertion and the multifactorial aspect of clinical outcomes in this challenging patient population.

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., 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 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: Groff. 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: Groff. Study supervision: Kaiser.

References

1. Andersen T, Christensen FB, Niedermann B, Helmig P, Høy K, Hansen ES, et al: Impact of instrumentation in lumbar spinal fusion in elderly patients: 71 patients followed for 2-7 years. **Acta Orthop 80:**445–450, 2009
2. Bjarke Christensen F, Stender Hansen E, Laursen M, Thomsen K, Bünger CE: Long-term functional outcome of pedicle screw instrumentation as a support for posterolateral spinal fusion: randomized clinical study with a 5-year follow-up. **Spine (Phila Pa 1976) 27:**1269–1277, 2002
3. Christensen FB, Laursen M, Gelineck J, Eiskjaer SP, Thomsen K, Bunger CE: Interobserver and intraobserver agreement of radiograph interpretation with and without pedicle screw implants: the need for a detailed classification system in posterolateral spinal fusion. **Spine (Phila Pa 1976) 26:**538–544, 2001
4. Epstein NE: An analysis of noninstrumented posterolateral lumbar fusions performed in predominantly geriatric patients using lamina autograft and beta tricalcium phosphate. **Spine J 8:**882–887, 2008
5. Fischgrund JS, Mackay M, Herkowitz HN, Brower R, Montgomery DM, Kurz LT: 1997 Volvo Award Winner in Clinical Studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. **Spine (Phila Pa 1976) 22:**2807–2812, 1997
6. Fritzell P, Hägg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study Group: Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish Lumbar Spine Study Group. **Spine (Phila Pa 1976) 27:**1131–1141, 2002
7. Jäger M, Seller K, Raab P, Krauspe R, Wild A: Clinical outcome in monosegmental fusion of degenerative lumbar instabilities: instrumented versus non-instrumented. **Med Sci Monit 9:**CR324–CR327, 2003
8. Korsgaard M, Christensen FB, Thomsen K, Hansen ES, Bünger C: The influence of lumbar lordosis on spinal fusion and functional outcome after posterolateral spinal fusion with and without pedicle screw instrumentation. **J Spinal Disord Tech 15:**187–192, 2002
9. Lorenz M, Zindrick M, Schwaegler P, Vrbos L, Collatz MA, Behal R, et al: A comparison of single-level fusions with and without hardware. **Spine (Phila Pa 1976) 16 (8 Suppl):**S455–S458, 1991
10. 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 12: pedicle screw fixation as an adjunct to posterolateral fusion for low-back pain. **J Neurosurg Spine 2:**700–706, 2005
11. Thomsen K, Christensen FB, Eiskjaer SP, Hansen ES, Fruensgaard S, Bünger CE: 1997 Volvo Award Winner in Clinical Studies. The effect of pedicle screw instrumentation on functional outcome and fusion rates in posterolateral lumbar spinal fusion: a prospective, randomized clinical study. **Spine (Phila Pa 1976) 22:**2813–2822, 1997
12. Tsutsumimoto T, Shimogata M, Yoshimura Y, Misawa H: Union versus nonunion after posterolateral lumbar fusion: a comparison of long-term surgical outcomes in patients with degenerative lumbar spondylolisthesis. **Eur Spine J 17:**1107–1112, 2008
13. Zdeblick TA: A prospective, randomized study of lumbar fusion. Preliminary results. **Spine (Phila Pa 1976) 18:**983–991, 1993

Manuscript submitted March 13, 2014.
Accepted April 8, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14277.
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.

J Neurosurg Spine 21:79–90, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 13: Injection therapies, low-back pain, and lumbar fusion

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

[1]Bone and Joint Clinic of Houston, Houston, Texas; [2]Department of Neurosurgery, University of Wisconsin, Madison, Wisconsin; [3]Center for Sports Medicine and Orthopaedics, Chattanooga, Tennessee; [4]Alan and Jacqueline Stuart Spine Research Center, Department of Neurosurgery, Lahey Clinic, Burlington, and Tufts University School of Medicine, Boston, Massachusetts; [5]Department of Neurological Surgery, University of California, San Francisco, California; [6]Department of Neurosurgery, University of Utah, Salt Lake City, Utah; [7]Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, New York, New York; [8]Department of Orthopaedic Surgery, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York; [9]Department of Neurosurgery, Brigham and Women's Hospital, Boston, Massachusetts; [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 medical literature continues to fail to support the use of lumbar epidural injections for long-term relief of chronic back pain without radiculopathy. There is limited support for the use of lumbar epidural injections for short-term relief in selected patients with chronic back pain. Lumbar intraarticular facet injections are not recommended for the treatment of chronic lower-back pain. The literature does suggest the use of lumbar medial nerve blocks for short-term relief of facet-mediated chronic lower-back pain without radiculopathy. Lumbar medial nerve ablation is suggested for 3–6 months of relief for chronic lower-back pain without radiculopathy. Diagnostic medial nerve blocks by the double-injection technique with an 80% improvement threshold are an option to predict a favorable response to medial nerve ablation for facet-mediated chronic lower-back pain without radiculopathy, but there is no evidence to support the use of diagnostic medial nerve blocks to predict the outcomes in these same patients with lumbar fusion. There is insufficient evidence to support or refute the use of trigger point injections for chronic lower-back pain without radiculopathy.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14281)

Key Words • **fusion** • **lumbar spine** • **epidural steroid injection** • **facet block** • **trigger point injection** • **low-back pain** • **practice guidelines**

Therapeutic Recommendations

There is no new evidence that conflicts with the previous recommendations regarding injection therapies published in the original version of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine."[27]

Lumbar Epidural Steroid Injections

Grade C

Lumbar epidural steroid injections (ESIs) are an option for the short-term relief of chronic low-back pain without radiculopathy in patients with degenerative disease of the lumbar spine (Level III evidence).

Caudal ESIs are an option for decreasing low-back pain of greater than 6 weeks' duration, without radiculopathy, in patients with degenerative disease of the lumbar spine (Level III evidence).

Lumbar Facet Injections

Grade B

Intraarticular injections of lumbar facet joints are not suggested for the treatment of facet-mediated chronic low-back pain without radiculopathy in cases of degenerative disease of the lumbar spine (single Level II study and single Level III study).

Lumbar medial nerve blocks are suggested for the

Abbreviations used in this paper: ESI = epidural steroid injection; NRS = Numerical Rating Scale; ODI = Oswestry Disability Index; RCT = randomized control trial; TPI = trigger point injection; VAS = visual analog scale.

short-term relief of facet-mediated chronic low-back pain without radiculopathy in patients with degenerative disease of the lumbar spine (single Level II study and single Level III study).

Lumbar medial nerve ablation is suggested for the short-term (3- to 6-month) relief of facet-mediated pain in patients who have chronic lower-back pain without radiculopathy from degenerative disease of the lumbar spine (4 Level II studies).

Lumbar Trigger Point Injections

Grade B

Trigger point injections (TPIs) performed as dry needling, with anesthetics alone or with steroids, are not recommended in patients with chronic low-back pain without radiculopathy from degenerative disease of the lumbar spine because a long-lasting benefit has not been demonstrated (Level II evidence).

Diagnostic Recommendations

Grade B

To establish the diagnosis of lumbar facet-mediated pain, the double-injection technique with an improvement threshold of 80% or greater is suggested (single Level I study).

Grade C

Diagnostic facet blocks by the double-injection technique with an improvement threshold of 80% are an option for predicting a favorable response to facet medial nerve ablation by thermocoagulation for facet-mediated chronic low-back pain without radiculopathy in patients with degenerative disease of the lumbar spine (single Level II study).

Grade I: Inconclusive

There is no evidence to support the use of diagnostic facet blocks as a predictor of lumbar fusion outcome in patients with chronic low-back pain from degenerative lumbar disease (conflicting Level IV evidence).

Rationale

Since the original publication of the Lumbar Fusion Guidelines, injection techniques using an anesthetic agent, typically in combination with a steroid, continue to be widely used in the treatment of patients with chronic low-back pain.[27] An updated analysis of the literature regarding these treatments was performed from July 2003, the termination point of the previous guidelines, through the end of 2011. As was the case in the original guidelines, an attempt was made to answer 3 questions:

1) Are lumbar ESIs effective for improving the outcomes of patients with chronic low-back pain resulting from degenerative disease of the lumbar spine?

2) Are lumbar facet injections effective for improving the outcomes of patients with chronic low-back pain resulting from degenerative disease of the lumbar spine?

3) Are lumbar TPIs effective for improving the outcomes of patients with chronic low-back pain resulting from degenerative disease of the lumbar spine?

Search Criteria

A computerized search of articles published from July 2003 through the year 2011 in the National Library of Medicine's MEDLINE database was conducted using the online search engine "PubMed." The search chain included the following terms: ("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields]) AND ("Injections, Spinal"[MeSH] OR "Injections, Intra-Articular"[MeSH] OR "Anesthesia, Epidural"[MeSH] OR "Nerve Block"[MeSH] OR trigger point injection [title] OR trigger point injections[title] OR (facet joint injection[title] OR facet joint injections[title] OR (epidural steroid injection[title] OR epidural steroid injections [title]) OR epidural steroid block[title] OR (caudal injection[title] OR caudal injections[title]) OR (caudal block[title] OR caudal blockade[title] OR caudal blocks [title]) OR (selective nerve root injection[title] OR selective nerve root injections[title]) OR (selective nerve root block[title] OR selective nerve root blocks[title]) OR (transforaminal injection[title] OR transforaminal injections[title] OR (transforaminal block[title] OR transforaminal blocks[title])) OR (block[title] OR block/activation[title] OR block/cytological[title] OR block/intra osseous[title] OR block/mylohyoid[title] OR block/neurolysis[title] OR block/sick[title] OR block/western[title] OR block'[title] OR block's[title] OR block98[title] OR blockable[title] OR blockad[title] OR blockada[title] OR blockade[title] OR blockade/myosin[title] OR block ade/thiazide[title] OR blockade'[title] OR blockaded [title] OR blockaden[title] OR blockader[title] OR block aders[title] OR blockaders/admin[title] OR blockades [title] OR blockading[title] OR blockador[title] OR block age[title] OR blockages[title] OR blockain[title] OR blockaine[title] OR blockal[title] OR blockase[title] OR blockboard[title] OR blockbuilding[title] OR block buster[title] OR blockbuster'[title] OR blockbusters[title] OR blockcourse[title] OR blockcycler[title] OR block dissection[title] OR blocke[title] OR blocked[title] OR blocked'[title] OR blocker[title] OR blocker/5[title] OR blocker/beta[title] OR blocker/calcium[title] OR blocker/carbonic[title] OR blocker/diuretic[title] OR blocker/drug [title] OR blocker/hydrochlorothiazide[title] OR blocker/statin[title] OR blocker/thiazide[title] OR blocker/vaso dilator[title] OR blocker's[title] OR blockerette[title] OR blockers[title] OR blockers/ace[title] OR blockers'[title] OR blockes[title] OR blockexcision[title] OR blockface [title] OR blockheads[title] OR blockholer[title] OR blocki[title] OR blockinducing[title] OR blockiness[title] OR blocking[title] OR blocking/deblocking[title] OR blocking/diuretic[title] OR blocking/percolation[title] OR blocking/unblocking[title] OR blocking'[title] OR block ings[title] OR blocklength[title] OR blockley[title] OR blockmakers[title] OR blockmaking[title] OR blockmilk [title] OR blockout[title] OR blockpnea[title] OR block polymer[title] OR blocks[title] OR blocks'[title] OR block sequences[title] OR blockset[title] OR blocksom[title] OR

blockwise[title] OR blockwriting[title] OR blocky[title] OR blockypnea[title] OR blockzone[title]) OR (facet joint block[All Fields] OR facet joint blocks[All Fields]) OR (median nerve block[title] OR median nerve block ade[title]) OR median nerve injection[title] OR (trigger point injection[title] OR trigger point injections[title]) OR (trigger[All Fields] AND (point block[title] OR point blocks[title])))) AND ((“Lumbosacral Region”[MeSH] OR “Lumbar Vertebrae”[MeSH]) OR lumbar[title]) AND ((“2003”[PDAT]: “3000”[PDAT]) AND “humans” [MeSH Terms] AND English[lang]).

The search was limited to English-language publications and human subjects. Nonsystematic reviews were discarded, but the bibliographies from these papers were searched for any additional relevant references. The search yielded 249 new references for this paper. Papers selected were confined to studies of chronic low-back pain (> 3–6 months) due to lumbar degenerative disease without deformity and without radiculopathy. The results of the search were divided into 3 categories depending on the type of injection investigated: ESIs, facet injections, and TPIs. All papers providing Level II or better evidence were included. In the absence of Level 1 or Level II data, Level III papers were included in the analysis. Papers with Level IV evidence were referenced in the discussion but not included in the evidentiary tables.

Scientific Foundation

Use of Lumbar ESIs (Interlaminar Injections, Caudal Injections, Transforaminal Injections) in the Treatment of Chronic Low-Back Pain Due to Degenerative Disease of the Lumbar Spine

Epidural injections continue to be used extensively in the treatment of spinal pain.[11,34] The evaluation of ESIs for chronic lower-back pain without radiculopathy remains minimal. In the previous review of this topic,[27] 4 randomized control trials (RCTs) were found to evaluate the effectiveness of epidural injections in the treatment of chronic lower-back pain.[4,9,28,30] All 4 of these studies were reported as RCTs but were greatly underpowered and represented equivalence trials without true control groups. By the criteria of the current report, these studies are classified as Level III data and give little support for the use of lumbar ESIs in chronic back pain for anything more than short-term relief (< 2 weeks). They are referenced in the bibliography but not in the evidence table (see Table 1). Since the completion of the previous review of this topic, a prospective cohort study published in 2004 by Buttermann evaluated 232 patients, age 18–65 years, with low-back pain of greater than 1 year in duration, in whom conservative maneuvers failed.[5] The patients were diagnosed with degenerative disc disease without stenosis or listhesis. They received 1–3 interlaminar or transforaminal steroid injections guided by fluoroscopy and were followed up for up to 2 years. Modic endplate changes on MRI, indicative of vertebral inflammation, were observed in 93 of the study participants. Buttermann predicted that these participants would appreciate more frequent relief of low-back pain after ESIs than would the group without inflammatory changes. Validated outcome measures were used, including the visual analog scale (VAS), Oswestry Disability Index (ODI), and pain drawings. Medication usage and the degree of patient satisfaction were also recorded. A subgroup of patients was randomized to receive a discogram with or without steroids. For patients with inflammatory endplate changes, 55% were satisfied with the degree of pain relief up to 3 months after the injection, although a clinically relevant improvement was not observed in the VAS or ODI scores. A similar finding was observed in the noninflammatory cohort with 47% satisfied. Improvement in both groups declined over time. While the baseline differences in ODI scores between the two groups was not different prior to treatment, comparison of these scores for the two groups at 3- and 6-month follow-up showed a statistically greater improvement for the group with inflammatory changes ($p < 0.001$), though neither group demonstrated a statistically significant improvement over baseline scores. This study has been cited as providing support for the short-term benefit of ESIs in decreasing chronic low-back pain, although there was no objective improvement in either group observed with the validated outcome measures. Furthermore, the conclusions of the study are severely compromised by the high dropout rate at final follow-up: 51% of the original patients in the inflammatory group and 60% of those in the noninflammatory cohort were lost to follow-up. In an equivalence trial, Manchikanti et al. published a randomized controlled and double-blinded study of 70 patients with lower-back pain and no radiculopathy or evidence of disc herniation on MRI.[19] Thirty-five of the patients were randomized to Group I in which the patients received interlaminar injections of anesthetic only and 35 were randomized to Group II in which the patients received interlaminar injections of an anesthetic and a steroid. Validated outcomes measures, including the ODI for functional assessment and the Numerical Rating Scale (NRS), were recorded at baseline and at 3, 6, and 12 months. Greater than a 50% improvement in pain or function from baseline was required for significance. Significant pain relief was recorded in 74% of Group I and 63% of Group II, while significant functional improvement was achieved in 71% of Group I and 60% of Group II. The overall average number of injections for the two groups over the year of follow-up was 4. This study, while suggestive, suffers from being an equivalence study without appropriate placebo control, from being underpowered, and from being a preliminary report.

Three systematic reviews were identified during the current search (Table 1). Abdi et al. performed a review of the literature from published 1966 to 2006 on cervical, thoracic and lumbar ESIs.[1] For the lumbar spine, 13 randomized control trials (RCTs) studies for transforaminal injections, and 8 RCTs and 5 prospective trials for caudal injections. The majority of these studies investigated the utility of these treatments for radiculopathy. With respect to chronic low-back pain, the Buttermann study, reviewed above, was felt to provide indeterminate evidence that ESIs were effective in managing chronic low-back pain when the transforaminal and interlaminar techniques were used.[5] In addition, the authors concluded there was

TABLE 1: Lumbar fusion injection procedures: epidural steroid injections

Authors & Year	Level of Evidence & Study Type	Description of Study*	Comment
Staal et al., 2009	I, systematic review	This is an update of the previous Cochrane Review on this topic (2000) w/ the literature search running 1999–2007 in patients 18–70 yrs of age. RCTs only were evaluated & 8 studies were added to the original review's 10 studies & thus conclusions are based on the combined data of 18 RCTs. Facet injections (intraarticular, periarticular, or median nerve), epidural, & local sites were studied. It was noted that in the period since the 1st review there had been no improvement in the quality of the clinical literature.	Based on this analysis, the authors concluded there is moderate evidence that epidural corticosteroid injections are not more effective than a placebo for pain relief; there is limited evidence that injections are not more effective than a placebo for general improvement & work disability & limited evidence they are not more effective than other drug treatments.
Buttermann, 2004	III, prospective comparative study	Cohort study of 232 patients age 18–65 yrs w/ >1 yr of lower-back pain only, w/ failure of conservative care (received 1–3 interlaminar or transforaminal injections w/ fluoroscopy 1 wk apart). 93 patients had Modic inflammatory changes & 139 did not. Outcome measures were VAS, ODI, & pain drawings. Also medication usage was monitored & patient opinion of success was recorded. A 2nd part of the trial was randomization of smaller subgroups to intradiscal steroid injections or not.	The authors noted a greater improvement in ODI score (1–3 & 4–5 mos of follow-up) & pain drawings in patients w/ Modic I changes (4–6 mos of follow-up) than in those w/o Modic I changes. Neither group, however, had a significant improvement in ODI vs baseline levels. Furthermore, the dropout rates in the Modic I groups were 51% over the 1- to 2-yr period & 60% in the other group, resulting in a down-classification of this study to Level III. The authors concluded that patients w/ >1 yr of axial back pain "may have short-term benefit from ESI" (25%–35%). The value of this study is severely limited by the high dropout rate.
Manchikanti et al., 2010[19]	III, RCT	Prospective RCT, double-blinded trial. 70 patients randomly assigned to 1 of 2 groups by computer-generated random allocation sequence: Group I, local anesthetic only; Group II, local anesthetic mixed w/ nonparticulate betamethasone. Outcome measures included the NRS, ODI 2.0, employment status, & opioid intake. Assessments done at baseline & 3, 6, & 12 mos posttreatment. Significant pain relief &/or improvement in disability were defined as ≥50% improvement.	Significant pain relief (≥50%) was demonstrated in 74% of patients in Group I & 63% in Group II. Functional status improvement (reduction of ≥50%) in the ODI scores was seen in 71% of Group I & 60% of Group II. Overall average no. of procedures/yr was approximately 4. The study is actually an equivalence study & did not contain an appropriate (placebo) control & was underpowered as well.

* Including analysis of methodological strengths/weaknesses.

moderate evidence in support of short- and long-term improvement in managing chronic low-back pain via the caudal approach. In 2009, Staal et al. published an update of a previous Cochrane Review, evaluating the literature from 1999 to March 2007 in patients 18–70 years of age.[32] Only RCTs involving facet, epidural, and local injections were considered. The authors noted that since their initial publication in 2001,[25] there was no apparent improvement in the quality of the evidence. With respect to ESIs for chronic low-back pain, the authors concluded that there was moderate evidence that ESIs are no more effective than a placebo for pain relief, that there is limited evidence ESIs and placebo are equally effective for general improvement in the work-disability population, and that there is limited evidence ESIs are more effective than other drug treatments. Parr et al. published a systematic review of studies published between 1966 and 2008 on lumbar interlaminar injections for the management of chronic low-back pain with and without radiculopathy.[26] They noted that the majority of these studies were done without fluoroscopic guidance. None of the RCTs identified investigated chronic low-back pain in the absence of a radiculopathy, and of the 30 observational studies, only the Buttermann article evaluated patients with isolated chronic low-back pain.[5] The authors concluded that the Buttermann article suggested some short-term but no long-term effect for ESIs on chronic lower-back pain.

Use of Lumbar Facet Injections for Chronic Low-Back Pain Due to Degenerative Disease of the Lumbar Spine

Lumbar facet (zygapophysial) joint injections have been used for both the diagnosis and treatment of facet-mediated low-back pain. Facet-mediated pain patterns have been explored by mapping the response to facet provocation and anesthesia injections in volunteers. These studies have yet to demonstrate a reliable pattern of pain produced by an injection within a particular lumbar facet joint.[7] When the data are combined from multiple studies, patterns emerge that suggest there is considerable overlap among all lumbar facet joints. Pain from the lower facet joints can be referred to the groin and deep posterior thigh, while the upper joints can lead to pain in the flank, hip, and upper lateral thigh. Pain referred below the knee is highly questionable. No physical or radiographic findings consistently correlate with the observations following facet blocks,[7] and the diagnosis of facet-mediated pain continues to rely on appropriately performed diagnostic facet blocks. The results of so-called double-block studies suggest that facet-mediated low-back pain is a cause of chronic pain in 9%–42% of patients with degenerative lumbar disease.[2,7,10,18,29]

Studies investigating the role of diagnostic facet joint blocks have been conducted in an attempt to improve the accuracy of this technique. Since the original guideline publication, a more uniform definition of a valid response has been adopted. It has been suggested that the double-block technique is the most reliable means of identifying facet-mediated pain, although this procedure is rarely performed during routine clinical practice. In the double-block technique, facet blocks are performed on two different dates with anesthetics that vary with respect to duration of the analgesic effect. A positive response requires that the patient's low-back pain significantly improve following both blocks for a period of time consistent with the anesthetic's duration of action.[2] To further refine the specificity of diagnostic facet injections, it has been suggested that the traditional threshold of greater than 50% pain relief be increased to greater than 80%. In a systematic review of 7 studies, Datta et al. presented Level I and II diagnostic evidence that the use of double controlled blocks and an 80% pain relief threshold produced the highest specificity in diagnosing facet-mediated back pain (Table 2).[10] They recommended that all future systematic reviews and investigations use these parameters as valid criteria to diagnose facet-mediated pain and evaluate the response to treatment. In an observational study, Manchikanti et al. demonstrated the improved sustainability of the diagnosis of lumbar facet-mediated pain at 2-year follow-up when comparing a group in which the 80% threshold was used for diagnosis and a group in which a 50% threshold for pain relief was used for diagnosis.[20] The diagnosis of facet-mediated pain was sustained in 89.5% of the patients diagnosed with the double-injection technique and an 80% threshold at 2 years versus only 51% of patients diagnosed with a double injection technique and a 50% threshold. The authors point out that utilizing the double-injection technique and an 80% threshold will diminish inappropriate and unnecessary treatment.[20]

Therapeutic facet blocks can be delivered in one of two manners: as an intraarticular injection into a facet joint or as a neural block of the medial nerve that innervates the facet capsule. In the previous review of this topic, 3 Level II studies addressed the efficacy of intraarticular injections in the facet joint.[6,17,23] Lilius et al. randomly assigned patients to one of three groups.[17] Group I received an intrafacet injection of steroid and anesthetic; Group II, pericapsular injections of steroid and anesthetic; and Group III, pericapsular injections of saline. The authors concluded that facet injections were a nonspecific form of treatment of lower-back pain that had good results depending more on psychosocial aspects of back pain. Carette et al. randomized 91 patients to facet injections of either methylprednisolone or saline.[6] No differences were seen between the groups at 1, 3, and 6 months postinjection. The authors concluded that injection of methylprednisolone into facet joints was of little treatment value. Marks et al. randomized 86 patients with chronic lower-back pain to receive either a facet injection with steroid and anesthetic or just an anesthetic block of the joint.[23] They concluded that at 3 months both types of injections were equally good diagnostically and equally unsatisfactory for treatment of chronic lower-back pain. The additional literature reviewed for the current report suggests that there is little evidence supporting the value of intraarticular facet blocks as a therapeutic option for chronic low-back pain, prompting one investigator to comment that the efficacy of these injections was no greater than a sham injection.[3] Datta et al.[10] performed a systematic, evidence-based review of the literature from 1966 through 2008 and identified 1438 articles investigating the utility of lumbar facet injections. They excluded

TABLE 2: Lumbar fusion injection procedures: facets*

Authors & Year	Level of Evidence & Study Type	Description of Study†	Comment
Datta et al., 2009	I, diagnostic; II–III, treatment	Systematic review (1966–2008) of diagnostic & therapeutic lumbar facet injections. Inclusion criteria included double-block technique & >80% pain relief for diagnostic blocks & decreased pain, improved function (along w/ return to work & decreased meds) for therapeutic injections. Treatment interventions included facet joint injections, median nerve blocks, & RF ablation. Short-term relief was defined as lasting ≤6 mos. Studies were ranked on the USPSTF modified levels of evidence scale (where Level III is opinion).	For diagnostic blocks, 7/1782 papers met inclusion criteria. These reported a prevalence of facet low-back pain of 16–40% (Level I). 1438 studies evaluated therapeutic facet intervention. For facet joint injections, the 5 RCTs & 15 observational studies did not meet inclusion criteria. Evidence for this intervention was Level III (opinion) & no recommendations could be made for its use. For facet joint nerve blocks, 2 RCTs met inclusion criteria. Based on these studies (Level II-1 evidence), the authors conclude there is strong evidence (moderate-quality evidence) for facet nerve blocks in short-term & long-term treatment of facet low-back pain. For RF ablation, there was 1 RCT & 2 observational studies that met inclusion criteria. The authors conclude there is Level II–III evidence for a treatment effect w/ a strong recommendation (low-quality evidence) for use.
Leclaire et al., 2001	I, RCT	Prospective double-blinded RCT of 70 patients w/ back pain for >3 mos. All responded to facet blocks & randomized to an RF ablation group or sham control group. Outcome measures were ODI & Roland-Morris instruments & VAS. At 4 wks, no difference in VAS or ODI scores; at 12 wks, no difference btwn the 2 groups on any outcome measures.	In patients selected by response to diagnostic facet injections, no difference in outcome measures at 4 & 12 wks.
Staal et al., 2009	II, systematic review of Level II data	Systematic review that is an update of the previous Cochrane Review on this topic (2001) w/ the literature search running 1999–2007 in patients 18–70 yrs of age. RCTs only were evaluated & 8 studies were added to the original review's 10 studies & thus conclusions are based on the combined data of 18 RCTs. Facet injections (intraarticular, periarticular, or median nerve), epidural, & local sites were studied. It was noted that in the period since the 1st review there had been no improvement in the quality of the clinical literature.	Based on this analysis, the authors concluded that there is moderate evidence that facet joint injections w/ steroids are no more effective than placebo injections for pain relief & improvement of disability. They further concluded there is insufficient evidence to support or refute the use of injection therapy for subacute & lower-back pain w/o radiculopathy regardless of type & dosage.
Nath et al., 2008	II, RCT	RCT looking at 40 patients (double-block technique w/ 80% pain reduction threshold led to 40/376 original patients being included). Patients randomized to active & control (no current) RF ablation. All were performed by the same operator; patients & treating doctors were blinded. Outcomes were global pain (10-point VAS scale) range of motion of back & hip. Follow-up was at 6 mos.	Global assessment was better for the active group ($p < 0.002$); generalized pain was improved more in the active group ($p < 0.02$); back pain was reduced more in the active group ($p < 0.004$); medication usage was decreased more in the active group ($p < 0.04$).
Manchikanti et al., 2010[21]	II, RCT	RCT included 120 patients w/ 60 patients in each group—local anesthetic alone or local anesthetic & steroids. Half of both groups also received Sarapin. Inclusion criteria were based on a positive response to diagnostic controlled, comparative local anesthetic lumbar facet joint blocks. Outcome measures included the NRS, ODI, opioid intake, & work status at baseline & 3, 6, 12, 18, & 24 mos.	Significant improvements, w/ significant pain relief of ≥50% & functional improvement of ≥40%, were observed in 85% of Group 1 & 90 of in Group II, at 2-yr follow-up. Patients experienced significant pain relief for 82–84 wks of 104 wks, requiring approximately 5–6 treatments w/ an average relief of 19 wks/episode of treatment.

(continued)

TABLE 2: Lumbar fusion injection procedures: facets* *(continued)*

Authors & Year	Level of Evidence & Study Type	Description of Study†	Comment
van Kleef et al., 1999	II, RCT	Prospective double-blinded RCT of 31 patients w/ >1 yr of low-back pain & positive response to facet blocks who were placed in an RF group or sham control group. Outcome measurements were VAS, perceived improvement, narcotics usage, & ODI. Differences in outcomes were statistically improved in the treatment group at 3, 6, & 12 mos.	In patients selected by response to diagnostic facet injections, those w/ RF ablation had superior outcomes at 3, 6, & 12 mos vs sham control.
Gallagher et al., 1994	II, RCT	Prospective RCT of 41 patients w/ chronic low-back pain who were selected based on a positive response to diagnostic facet blocks. Outcomes measures were the VAS and MPQ. Outcomes in patients receiving RF ablation were statistically superior to those in a placebo control group at 1 & 6 mos.	In patients selected by response to diagnostic facet injections, those w/ RF ablation had superior outcomes at 1 & 6 mos vs placebo control.
Marks et al., 1992	II, RCT	86 patients w/ chronic low-back pain were randomized to a facet block only w/ anesthetic or intraarticular injection w/ steroid & anesthetic. Immediate response was the same for both groups; steroid group was marginally better in a pain measure at 1 mo ($p < 0.05$), & by 3 mos only 2 patients reported any pain relief.	Facet blocks or injections w/ anesthetic & steroid were equally good diagnostically. Both were equally unsatisfactory treatment for chronic back pain.
Carette et al., 1991	II, RCT	91 patients w/ chronic lower-back pain who responded to facet joint injections w/ a diminution in pain were randomized to injections w/ steroid or injections w/ saline control. No difference in outcome measures at 1 & 3 mos btwn the 2 groups nor was there difference in sustained improvement from Mo 1 to Mo 6.	Injection of methylprednisolone into the facet joints is of little treatment value.
Lilius et al., 1990	II, RCT	109 patients were randomly assigned to 1 of 3 groups: Group 1, intrafacet injections of steroid & anesthetic; Group 2 pericapsular injections of steroid & anesthetic; & Group 3, pericapsular injection of saline as a control. Outcome measures were subjective pain scale, work, & disability income. Signs of inappropriate behavior prior to injection were best predictor of outcomes, & no difference in overall outcomes were found among the 3 groups.	Facet injections are a nonspecific form of treatment for low-back pain & good results depend on a tendency toward spontaneous regression of back pain & on psychosocial aspects of back pain.
Manchikanti et al., 2008	III, RCT	Double-blind RCT comparing 60 patients in Group I receiving local anesthetic only and 60 patients in Group II receiving steroids also diagnosed by double blocks with 80% relief threshold. VAS, ODI, opioid use, & RTW measured at baseline & 3, 6, & 12 mos. Intent-to-treat analysis was performed.	There was a significant improvement in VAS from baseline at all follow-up dates & there was no significant difference btwn groups (that is, no additional effect for steroids). 82% of patients showed improved pain & function for approximately 15 wks, w/ repeat injections required 3–4 times to maintain improvement for 1 yr. The study is criticized for lack of a placebo control.

* meds = medications; MPQ = McGill Pain Questionnaire; RF = radiofrequency; USPSTF = US Preventive Services Task Force.
† Including analysis of methodological strengths/weaknesses.

studies not evaluating patients with chronic low-back pain of more than 3 months' duration that was diagnosed as facet-mediated pain by the double-injection technique, with a greater than 80% pain relief threshold. Six RCTs and 15 observational studies were identified that evaluated the effectiveness of lumbar intraarticular facet injections. These studies were rejected due to poor methodology and failure to use the double-injection technique to confirm the diagnosis. Based on this systematic review of low-quality evidence, the authors concluded that there was no role for intraarticular facet injections as a treatment modality. This conclusion was supported by the update of the Cochrane Review published by Staal et al. in 2009.[32] These authors identified moderate-level evidence that facet joint injections with steroids are no more effective than placebo injections for relief of pain and disability.

The evidence for therapeutic efficacy is better for medial nerve blocks of the lumbar facet joint. In their systematic review, by Datta et al. also evaluated the role of lumbar facet nerve blocks as a therapeutic intervention.[10] They identified two RCTs that met inclusion criteria but no observational studies. Manchikanti et al. performed a double-blinded RCT of 120 patients with facet-mediated low-back pain of greater than 6 months' duration diagnosed using the double-injection technique and an 80% relief threshold.[22] All patients underwent a fluoroscopically guided injection of the medial nerve. Group I (n = 60) received anesthetic only and Group II (n = 60) received anesthetic and steroid. Half of each group also received Sarapin in the injectant. Multiple injections were performed at the discretion of the treating physician over 1 year. Validated outcome measures including the VAS and ODI were used along with nonvalidated measures of drug usage and return-to-work status. An intent-to-treat analysis was used to evaluate the data at final follow-up. Patients received up to 5 injections over the 1-year period with an average of 3.4 injections per patient. Improved pain scores, with over 50% pain relief reported in over 80% of the participants, were observed at 3, 6, and 12 months after the first injection when compared with baseline; however, no differences were observed between treatment groups. The ODI results were also significantly improved at 3, 6, and 12 months in all groups but with no differences between treatment groups. There was no significant decrease in opioid use observed in any group. These results support the premise that patients may experience significant pain relief from multiple injections for up to 44–45 weeks, with each injection providing on average of 15 weeks of pain relief for low-back pain and increased function as measured by the ODI. This study, an equivalence study that did not include a placebo control, provides moderate evidence that medial nerve injections confer short-term relief of chronic facet-mediated low-back pain. In a 2-year follow-up study of this same group of patients, Manchikanti et al. demonstrated that outcomes were sustained in both groups.[21] Pain relief of greater than 50% and functional improvement of greater than 40% were seen in 85% of Group I and 90% of Group II at 18 and 24 months. Continued need for repeated injections, with an average of 5 or 6 injections over the study period and duration of effect of 19 weeks, was seen in the longer follow-up.[21]

Ablation of the medial nerve, through radiofrequency thermocoagulation, is a variant of the facet nerve block. In the previous review of this topic, several papers were found testing the ability of facet blocks to predict outcomes from radiofrequency thermocoagulation.[27] Gallagher et al. performed a prospective, double-blinded RCT on 41 patients who reported either a strong or equivocal response to diagnostic facet blocks.[13] These 41 patients received either radiofrequency ablation with an anesthetic or just an anesthetic injection. Outcomes were assessed using the McGill Pain Questionnaire and VAS at 1 and 6 months. Patients who were strongly positive on facet blocks and received radiofrequency ablation did statistically better on both outcome measures at both times than those who were poor responders to facet blocks and received ablation. Van Kleef et al. randomized 31 patients who had responded strongly to facet blocks into two groups: one received radiofrequency ablation and the other received a sham control.[33] Both patients and treating doctors were blinded as to treatment who was in the control group. Outcomes were assessed using the VAS and ODI and by quantification of the amount of narcotic used. Outcomes were statistically superior in the radiofrequency group over the control at 3, 6, and 12 months. In a larger blinded RCT of 70 patients who had responded to facet blocks, Leclaire et al. measured outcomes after radiofrequency ablation using the VAS, ODI, and Roland-Morris disability questionnaire and found that results were superior only at 2 weeks, indicating no superiority for radiofrequency ablation for long-term relief of lower-back pain in this study.[16] Nath et al. conducted a randomized, double-blinded study of patients with chronic low-back pain of 2 years' duration in whom conservative treatment failed.[24] They included only patients with facet-mediated low-back pain, diagnosed by the double-block technique and a threshold of greater than 80% pain reduction. From a potential population of 376 candidates, only 40 patients fulfilled all the diagnostic criteria. These patients were randomized into a treatment group (n = 20), receiving active radiofrequency ablation, and a placebo group (n = 20), undergoing an identical sham procedure. Primary outcome measures included a VAS pain scale and a nonvalidated, self-reported 1- to 6-point global improvement scale. Lumbar range of motion and a 6-point quality of life scale were used as secondary outcome measures. Generalized pain, low-back pain, and referred pain were all significantly reduced in the treatment group compared with the control group at 6 months' follow-up. Although this is an underpowered study, the strict diagnostic inclusion criteria lend strength to its conclusions. This paper provides moderate evidence for the effectiveness of facet radiofrequency ablation in the short-term treatment of facet-mediated back pain.

Despite the increased diagnostic rigor seen more frequently in the newer literature (the double blocks and the 80% threshold for pain reduction), no new studies have appeared to suggest that diagnostic facet blocks can effectively predict the outcomes of surgical fusion in patients with chronic low-back pain from lumbar degenerative disease.

Use of Local Lumbar Injections (TPIs) in the Treatment of Chronic Low-Back Pain Due to Degenerative Disease of the Lumbar Spine

In the previous review of TPIs for chronic low-back pain, 4 Level II RCTs of small patient numbers were presented.[8,14,15,31] In a very small study, Hameroff et al. randomized, in a double-blind fashion, 15 patients into 3 groups: Group 1 received bupivacaine TPIs, Groups 2 received etidocaine injections, and Group 3 received a saline control injection.[15] Subjective reports of pain were obtained at 15 minutes, 1 day, and 7 days after injection. Trigger point injections with anesthetic were more effective than those with saline. Sonne et al.[31] prospectively randomized 30 patients with at least 1 month of lower-back pain into 2 groups in a double-blinded study: Group I received an injection of methylprednisolone with lignocaine and Group II received an injection of isotonic saline. Outcome measures were the VAS and lumbar range of motion. Significant decreases in VAS scores were seen in the anesthesia/steroid group while there was no difference between the two groups in terms of range of motion. Garvey et al. performed a randomized, double-blind evaluation of 63 patients with low-back pain unresponsive to 4 weeks of conservative care.[14] He divided the patients into 4 groups: Group I was treated with lidocaine TPIs, Group II with lidocaine and steroid TPIs, Group III with dry needling, and Group IV with acupressure and vapocoolant. More patients reported decreased pain in response to acupressure and coolant (63%) than to drug TPIs (42%), but the difference was not significant. The authors concluded that TPIs have some potential value in treating lower-back pain but that injecting a drug was not necessary. Collée et al., in a double-blind study, randomly assigned 41 patients to receive TPIs with 0.5% lignocaine or saline.[8] Outcome measures were the VAS and a pain-intensity scores measured 2 weeks after injection. The group receiving the anesthetic had a significantly better decrease in pain than did the saline group. For all of these studies, it should be noted that none of the patient groups fulfilled a definition of chronic lower-back pain (> 3 months' duration). In reviewing the literature for the current review, no high-quality studies on the efficacy of TPIs were found since the original Guideline publication. There have, however, been 2 published systematic reviews that focused partially or completely on TPIs (see Table 3).

In 2005 Furlan et al. published a Cochrane Review focusing on acupuncture and dry-needling for both acute and chronic low-back pain and reviewed the literature from 1996 to February 2003.[12] While 35 RCTs were identified, only 20 of these were in English and all of the RCTs were felt to have significant methodological flaws. With respect to dry needling for chronic low-back pain, the authors concluded that the evidence was insufficient and of exceedingly poor quality to formulate any meaningful recommendations. A more contemporary Cochrane Review of injection therapy for subacute and chronic low-back pain by Staal et al. included TPIs as a treatment alternative for chronic low-back pain patients.[32] The literature published between 1999 and 2007 was reviewed. The authors concluded, based on limited data, that TPIs with steroids are no more effective than placebo injections for pain relief and improvement of disability. They stated that there was insufficient evidence to support the use of injection therapy for subacute and chronic low-back pain without radiculopathy regardless of type and dosage. The studies reviewed in the original Guidelines as well as these 2 systematic reviews suggest no significant differences in treatment effect exist among the uses of an anesthetic, an anesthetic and steroid, or dry needling with TPIs. Any improvement seen with these techniques was only apparent in acute cases of low-back pain. No evidence was available to support the effectiveness of TPIs in the treatment of chronic low-back pain.

Summary

Based on the literature reviewed for the original guideline publication as well as this updated review, there is weak evidence that ESIs provide short-term relief of pain in patients with chronic low-back pain from degenerative lumbar disease. There is evidence that caudal ESIs are an option for decreasing pain for greater than 6 weeks in patients with chronic low-back pain from degenerative lumbar disease (Level III evidence).

Based on the original guidelines as well as this updated review, there is moderate evidence to recommend that the diagnosis of facet-mediated back pain be made with the double-injection technique and a greater than 80% improvement threshold (Level II evidence). There is moderate evidence supporting a recommendation that diagnostic facet blocks be used to predict a good response to facet medial nerve ablation by thermocoagulation for facet-mediated chronic low-back pain (Level II evidence). There is moderate evidence suggesting that there is no role for intraarticular facet injections in the treatment of chronic low-back pain from lumbar degenerative disease (Level II evidence against). There is moderate evidence supporting the use of facet medial nerve blocks to achieve short-term pain relief for patients with facet-mediated chronic low-back pain from degenerative lumbar disease (Level II evidence). There is moderate evidence that facet medial nerve ablation produces a short-term decrease (3–6 months) of facet-mediated chronic low-back pain (Level II evidence).

There is no evidence to support a recommendation that diagnostic blocks are useful predictors of surgical outcomes following lumbar fusion.

Based on the original guidelines as well as this updated literature review, there is no evidence to support the use of TPIs with a dry-needling technique, with anesthetics alone or accompanied by steroids, in the management of patients suffering from chronic low-back pain secondary to degenerative lumbar disease (Level IV evidence).

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.;

TABLE 3: Lumbar fusion injection procedures: trigger point injections

Authors & Year	Level of Evidence & Study Type	Description of Study*	Comment
Staal et al., 2009	II, systematic review of Level II data	This is an update of the previous Cochrane Review on this topic (2000) w/ the literature search running 1999–2007 in patients 18–70 yrs of age. RCTs only were evaluated and 8 studies were added to the original review's 10 studies & thus conclusions are based on the combined data of 18 RCTs. Facet injections (intraarticular, periarticular, or median nerve), epidural, & local sites were studied. It was noted that in the period since the 1st review there had been no improvement in the quality of the clinical literature.	The authors concluded that there is moderate and limited evidence that TPIs w/ steroids are not more effective than placebo injections for pain relief and improvement of disability. They further concluded there is insufficient evidence to support or refute the use of injection therapy for subacute and lower-back pain w/o radiculopathy regardless of type & dosage.
Furlan et al., 2005	II, systematic review	This is an update of a previous Cochrane Review on the same subject & runs from 1996–2003. This included 35 RCTs, only 20 of which were published in English. This English database was covered in *J Neurosurg Spine* 2:707–715 and adds nothing new.	The authors concluded that for dry needling (TPIs) in chronic lower-back pain, no clear recommendations can be made because of studies of small sample sizes & low methodological quality.
Collée et al., 1991	II, RCT	RCT in which 41 patients were randomized to an injection w/ 0.5% lignocaine or an equal amount of saline. Follow-up was at 2 wks w/ VAS & subjective pain scores. Significantly more patients were improved after lignocaine (52%) than after saline injections (39%) ($p < 0.05$). Subgroup analysis found the effect only applied to patients treated in a rheumatology group & not in a family practice group.	TPI w/ anesthetic is more effective than saline in reducing back pain if administered by someone familiar & competent w/ technique.
Garvey et al., 1989	II, RCT	RCT in which 63 patients w/ low-back pain were treated for >4 wks & did not improve. Randomly assigned to 1 of 4 groups: Group 1, TPI of lidocaine; Group 2, injection of lidocaine w/ steroid; Group 3, dry needling of the TPI; & Group 4, vapocoolant spray w/ acupressure. Dry needle group responded at least as well as the groups w/ medication at injection (63% vs 42%; $p < 0.09$).	TPIs decrease low-back pain, but it does not appear necessary to inject any drug to achieve this.
Sonne et al., 1985	II, RCT	RCT in which 31 patients were randomly assigned to 1 of 2 injection groups: Group 1, methylprednisolone w/ lignocaine; Group 2, isotonic saline injection. Outcomes measured were VAS, self-assessment, & spinal range of motion. VAS and self-assessment scores were significantly better for the steroid injection, & no change in spinal range of motion was seen in either group.	Injection of steroid & anesthetic was significantly better for short-term relief of low-back pain than saline injection.
Hameroff et al., 1981	II, RCT	RCT in which 15 patients w/ low-back pain were randomly assigned to 1 of 3 groups in a double-blind, crossover study: Group 1, injected w/ bupivacaine; Group 2, injected w/ etidocaine; & Group 3, injected w/ saline as a control. Outcomes were measured according to subjective responses to 6 pain-related categories. Injection w/ anesthetic was superior to that w/ saline.	Injection of an anesthetic was preferred to a placebo injection of saline.

* Including analysis of methodological strengths/weaknesses.

Kevin Cockroft, 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 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: Watters. 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: Watters. Study supervision: Kaiser.

References

1. Abdi S, Datta S, Trescot AM, Schultz DM, Adlaka R, Atluri SL, et al: Epidural steroids in the management of chronic spinal pain: a systematic review. **Pain Physician 10:**185–212, 2007
2. Bogduk N: Evidence-informed management of chronic low back pain with facet injections and radiofrequency neurotomy. **Spine J 8:**56–64, 2008
3. Bogduk N: A narrative review of intra-articular corticosteroid injections for low back pain. **Pain Med 6:**287–296, 2005
4. Breivik H, Helsa PE, Molnar I, Lind B: Treatment of chronic low back pain and sciatica. Comparison of caudal injections of bupivacaine and methylprednisolone with bupivacaine followed by saline. **Adv Pain Res Ther 1:**927–932, 1976
5. Buttermann GR: Treatment of lumbar disc herniation: epidural steroid injection compared with discectomy. A prospective, randomized study. **J Bone Joint Surg Am 86-A:**670–679, 2004
6. Carette S, Marcoux S, Truchon R, Grondin C, Gagnon J, Allard Y, et al: A controlled trial of corticosteroid injections into facet joints for chronic low back pain. **N Engl J Med 325:**1002–1007, 1991
7. Cohen SP, Raja SN: Pathogenesis, diagnosis, and treatment of lumbar zygapophysial (facet) joint pain. **Anesthesiology 106:** 591–614, 2007
8. Collée G, Dijkmans BA, Vandenbroucke JP, Cats A: Iliac crest pain syndrome in low back pain. A double blind, randomized study of local injection therapy. **J Rheumatol 18:**1060–1063, 1991
9. Dallas TL, Lin RL, Wu WH, Wolskee P: Epidural morphine and methylprednisolone for low-back pain. **Anesthesiology 67:**408–411, 1987
10. Datta S, Lee M, Falco FJ, Bryce DA, Hayek SM: Systematic assessment of diagnostic accuracy and therapeutic utility of lumbar facet joint interventions. **Pain Physician 12:**437–460, 2009
11. DePalma MJ, Slipman CW: Evidence-informed management of chronic low back pain with epidural steroid injections. **Spine J 8:**45–55, 2008
12. Furlan AD, van Tulder M, Cherkin D, Tsukayama H, Lao L, Koes B, et al: Acupuncture and dry-needling for low back pain: an updated systematic review within the framework of the Cochrane collaboration. **Spine (Phila Pa 1976) 30:**944–963, 2005
13. Gallagher J, Di Vadi PLP, Wedley JR, Hamann W, Ryan P, Chikanza I, et al: Radiofrequency facet joint denervation in the treatment of low back pain: a prospective controlled double-blind study to assess its efficacy. **Pain Clin 7:**193–198, 1994
14. Garvey TA, Marks MR, Wiesel SW: A prospective, randomized, double-blind evaluation of trigger-point injection therapy for low-back pain. **Spine (Phila Pa 1976) 14:**962–964, 1989
15. Hameroff SR Sr, Crago BR, Blitt CD, Womble J, Kanel J: Comparison of bupivacaine, etidocaine, and saline for trigger-point therapy. **Anesth Analg 60:**752–755, 1981
16. Leclaire R, Fortin L, Lambert R, Bergeron YM, Rossignol M: Radiofrequency facet joint denervation in the treatment of low back pain: a placebo-controlled clinical trial to assess efficacy. **Spine (Phila Pa 1976) 26:**1411–1417, 2001
17. Lilius G, Harilainen A, Laasonen EM, Myllynen P: Chronic unilateral low-back pain. Predictors of outcome of facet joint injections. **Spine (Phila Pa 1976) 15:**780–782, 1990
18. Manchikanti L, Boswell MV, Singh V, Pampati V, Damron KS, Beyer CD: Prevalence of facet joint pain in chronic spinal pain of cervical, thoracic, and lumbar regions. **BMC Musculoskelet Disord 5:**15, 2004
19. Manchikanti L, Cash KA, McManus CD, Pampati V, Benyamin RM: Preliminary results of a randomized, double-blind, controlled trial of fluoroscopic lumbar interlaminar epidural injections in managing chronic lumbar discogenic pain without disc herniation or radiculitis. **Pain Physician 13:**E279–E292, 2010
20. Manchikanti L, Pampati S, Cash KA: Making sense of the accuracy of diagnostic lumbar facet joint nerve blocks: an assessment of the implications of 50% relief, 80% relief, single block, or controlled diagnostic blocks. **Pain Physician 13:**133–143, 2010
21. Manchikanti L, Singh V, Falco FJ, Cash KA, Pampati V: Evaluation of lumbar facet joint nerve blocks in managing chronic low back pain: a randomized, double-blind, controlled trial with a 2-year follow-up. **Int J Med Sci 7:**124–135, 2010
22. Manchikanti L, Singh V, Falco FJE, Cash KA, Pampati V: Lumbar facet joint nerve blocks in managing chronic facet joint pain: one-year follow-up of a randomized, double-blind controlled trial: Clinical Trial NCT00355914. **Pain Physician 11:**121–132, 2008
23. Marks RC, Houston T, Thulbourne T: Facet joint injection and facet nerve block: a randomised comparison in 86 patients with chronic low back pain. **Pain 49:**325–328, 1992
24. Nath S, Nath CA, Pettersson K: Percutaneous lumbar zygapophysial (facet) joint neurotomy using radiofrequency current, in the management of chronic low back pain: a randomized double-blind trial. **Spine (Phila Pa 1976) 33:**1291–1298, 2008
25. Nelemans PJ, deBie RA, deVet HC, Sturmans F: Injection therapy for subacute and chronic benign low back pain. **Spine (Phila Pa 1976) 26:**501–515, 2001
26. Parr AT, Diwan S, Abdi S: Lumbar interlaminar epidural injections in managing chronic low back and lower extremity pain: a systematic review. **Pain Physician 12:**163–188, 2009
27. 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 13:

injection therapies, low-back pain, and lumbar fusion. **J Neurosurg Spine 2:**707–715, 2005
28. Rocco AG, Frank E, Kaul AF, Lipson SJ, Gallo JP: Epidural steroids, epidural morphine and epidural steroids combined with morphine in the treatment of post-laminectomy syndrome. **Pain 36:**297–303, 1989
29. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N: Clinical features of patients with pain stemming from the lumbar zygapophysial joints. Is the lumbar facet syndrome a clinical entity? **Spine (Phila Pa 1976) 19:**1132–1137, 1994
30. Serrao JM, Marks RL, Morley SJ, Goodchild CS: Intrathecal midazolam for the treatment of chronic mechanical low back pain: a controlled comparison with epidural steroid in a pilot study. **Pain 48:**5–12, 1992
31. Sonne M, Christensen K, Hansen SE, Jensen EM: Injection of steroids and local anaesthetics as therapy for low-back pain. **Scand J Rheumatol 14:**343–345, 1985
32. Staal JB, de Bie RA, de Vet HC, Hildebrandt J, Nelemans P: Injection therapy for subacute and chronic low back pain: an updated Cochrane review. **Spine (Phila Pa 1976) 34:**49–59, 2009
33. van Kleef M, Barendse GA, Kessels A, Voets HM, Weber WE, de Lange S: Randomized trial of radiofrequency lumbar facet denervation for chronic low back pain. **Spine (Phila Pa 1976) 24:**1937–1942, 1999
34. Young IA, Hyman GS, Packia-Raj LN, Cole AJ: The use of lumbar epidural/transforaminal steroids for managing spinal disease. **J Am Acad Orthop Surg 15:**228–238, 2007

Manuscript submitted March 20, 2014.
Accepted April 9, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14281.

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.

J Neurosurg Spine 21:91–101, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 14: Brace therapy as an adjunct to or substitute for lumbar fusion

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

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

The utilization of orthotic devices for lumbar degenerative disease has been justified from both a prognostic and therapeutic perspective. As a prognostic tool, bracing is applied prior to surgery to determine if immobilization of the spine leads to symptomatic relief and thus justify the performance of a fusion. Since bracing does not eliminate motion, the validity of this assumption is questionable. Only one low-level study has investigated the predictive value of bracing prior to surgery. No correlation between response to bracing and fusion outcome was observed; therefore a trial of preoperative bracing is not recommended. Based on low-level evidence, the use of bracing is not recommended for the prevention of low-back pain in a general working population, since the incidence of low-back pain and impact on productivity were not reduced. However, in laborers with a history of back pain, a positive impact on lost workdays was observed when bracing was applied. Bracing is recommended as an option for treatment of subacute low-back pain, as several higher-level studies have demonstrated an improvement in pain scores and function. The use of bracing following instrumented posterolateral fusion, however, is not recommended, since equivalent outcomes have been demonstrated with or without the application of a brace.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14282)

Key Words • **brace** • **bracing** • **low-back pain** • **lumbar fusion** • **practice guidelines** • **spine**

Recommendations

There is no evidence that conflicts with the previous recommendations published in the original version of the guidelines for the use of lumbar bracing in the treatment of low-back pain.

Grade B

The prescription of a lumbar brace is useful for the secondary prevention of low-back pain by reducing the number of days of self-reported low-back pain and days lost to work in laborers with a history of low-back pain (single Level I study and multiple Level II studies).

For primary prevention, the use of a lumbar corset does not prevent the development of low-back pain in the general working population (multiple Level II studies).

For patients presenting with low-back pain, the prescription of a lumbar support in the setting of subacute pain (< 6 months' duration) reduced the visual analog scale (VAS) pain score and medication usage and im-

Abbreviations used in this paper: DPQ = Dallas Pain Questionnaire; ODI = Oswestry Disability Index; PLF = posterolateral lumbar fusion; RMDQ = Roland-Morris Disability Questionnaire; RSA = roentgen stereophotogrammetric analysis; SF-12 = 12-Item Short Form Health Survey; SF-36 = 36-Item Short Form Health Survey; TEPF = temporary external pedicle fixation; VAS = visual analog scale.

proved functional disability at 30–90 days (single Level I study and multiple Level II studies).

Grade C

The use of a brace following instrumented posterolateral lumbar fusion (PLF) for lumbar spondylosis is not supported due to equivalent outcomes with and without bracing (single Level II study).

Finally, a trial of preoperative bracing is not predictive of outcome for lumbar fusion in the setting of low-back pain (Level III evidence).

Rationale

Lumbosacral orthotics have been used for the prevention and treatment of a wide variety of degenerative disorders of the lumbar spine.[10,24,33] In addition, they have been used to improve outcome following lumbar fusion surgery and to aid in the selection of appropriate surgical candidates.[8] The potential mechanisms of action remain an area of debate and include limiting spinal range of motion, correcting posture and deformity, preventing gross trunk motion, increasing intraabdominal pressure, reducing force exerted by trunk muscles, providing soft-tissue massage and heat, and improving spinal proprioception.[5–7,19,20,24,33] Critics of lumbar supports have argued that bracing may provide workers with a false sense of support or allow muscles to atrophy, thereby increasing the potential for injury, particularly on discontinuation of use.[22,27] The clinical utility of lumbar bracing in the prevention and treatment of low-back pain remains controversial without conclusive evidence to support or refute the use of these devices.[13,18]

Braces have also been used in preoperative evaluation in an attempt to predict outcome following fusion surgery and used following lumbar surgery to promote a successful arthrodesis.[8,14] Because lumbar orthoses do not eliminate motion in the lumbar spine, their utility has been questioned.[2,3] The purpose of this review is to examine the medical evidence investigating the utility of brace therapy as strategy for prevention of low-back pain in the workplace, as a treatment for low-back pain, as a predictor of outcome following lumbar fusion surgery, and as an adjunct to lumbar fusion procedures.

Search Criteria

A computerized search of the National Library of Medicine database of the literature published from 2003 to 2011 was conducted using the following search terms: (“Lumbosacral Region [MeSH] OR “lumbar vertebrae [MeSH] or lumbar [title] or lumbosacral [title] AND (“low back pain [MeSH] OR “low back pain” [All Fields] OR “lower back pain” [All Fields] AND (“Orthotic Devices” [MeSH] OR “Braces” [MeSH] OR “brace” [title] OR “bracing” [title] OR “braces” [title]) AND ((“2003”[PDAT]: “3000” [PDAT]) AND “humans” [MeSH] AND English [lang]). After duplicates were discarded, 97 papers were identified, and their abstracts were reviewed. Eight relevant studies were identified and reviewed in detail, in addition to the 19 relevant studies from the previous guidelines.[28] In our previous guidelines, regarding the use of bracing and external fixation for fusion, we identified 10 relevant studies using temporary external pedicle fixation (TEPF) to predict the response to fusion for low-back pain. Because of a significant complication rate (20%–25%) and the uncertainty of TEPF to predict outcome following lumbar fusion, TEPF was not recommended as a screening modality for patients suffering with low-back pain. It is not considered a routine modality, and further discussion was eliminated from this review. Several review papers, meta-analyses, biomechanical studies, technical notes, and small case series served to provide supporting data. The bibliography of each paper was reviewed and other relevant studies were identified. All clinical studies providing Level III medical evidence or better regarding the use of lumbar brace therapy for the prevention and treatment of low-back pain, for the prediction of outcome following lumbar fusion surgery, and as an adjunct to fusion surgery are summarized in Tables 1–4.

Scientific Foundation

Bracing for Prevention of Low-Back Pain

Lumbar braces have been used as a means of preventing either initial (primary prevention) or recurrent (secondary prevention) episodes of low-back pain in industrial workers.[13,33] Van Poppel et al. randomized 282 individuals employed as baggage handlers into 4 groups: 1) education and lumbar brace, 2) education, 3) lumbar brace, and 4) no intervention.[34] Employees in Groups 1 and 3 wore soft lumbar braces for a 6-month period while working. For the entire cohort, there was no decrease in the incidence of reported back pain (36% for braced individuals and 34% for nonbraced) or in the number of workdays lost when comparing braced with nonbraced workers. A subgroup analysis of workers with a history of back pain revealed that the use of a soft lumbar brace reduced the number of days lost due to back pain from 6.5 to 1.2 days per month ($p = 0.03$). It should be noted that only 43% of the workers complied with the bracing protocol. Within the bracing cohort, there was no difference in the incidence of low-back pain or number of sick days among workers who complied and those who did not comply with the bracing protocol. The authors concluded that brace therapy does not diminish the incidence of low-back pain or time lost from work when used as a preventive strategy. The use of a lumbar support by workers with a previous history of low-back injury may reduce days lost due to low-back pain. Because of the high number of noncompliant workers, this study is considered to provide Level II medical evidence.

Reddell and colleagues randomized 642 individuals employed as baggage handlers into 4 groups: 1) education, 2) weightlifting belt–type brace, 3) education and brace, and 4) no intervention.[27] During an 8-month period, the authors examined the total incidence of reported low-back injury, lost or restricted workdays due to low-back pain, and Workers’ Compensation claims related to low-back pain. They found no differences among the

TABLE 1: Bracing for prevention of low-back pain: summary of evidence*

Authors & Year	Level of Evidence	Description	Comment
Kraus et al., 2002	II	Randomized trial of NYC home health attendants. 12,772 workers were randomized to 1 of 3 groups: 1) lumbar support, 2) safety meeting w/ information, or 3) no intervention. The outcome measure was self-reported back injury rates over a period of up to 28 mos, though it is uncertain how many participants reached the final time point. There are little data regarding participant demographics.	There was a trend toward fewer episodes of back injury in the lumbar support group compared w/ the information group, although the difference was not significant. The lumbar support group had fewer episodes than the no-intervention group. Compliance rates for the lumbar support group are difficult to determine, though a 97% compliance rate is mentioned.
van Poppel et al., 1998	II	282 (of 312) Dutch baggage handlers were randomized to 4 groups: 1) education & lumbar support, 2) education, 3) lumbar support, or 4) no intervention. Lumbar supports were used during work hours for a 6-mo period. Only a 43% compliance rate w/ soft brace. The use of a brace did not significantly decrease incidence of back pain (36% w/ brace vs 34% w/o brace) or the number of lost workdays. In 1 subgroup of pts (those w/ previous back pain) bracing reduced workdays lost to back pain from 6.5 to 1.2.	Lumbar supports do not reduce LBP incidence or sick leave when used as a preventive strategy for LBP. Unlike previous studies, this study showed no increase in the incidence of LBP in the groups discontinuing use of the belt.
Alexander et al., 1995	II	60 health care workers were divided into 2 groups: 1) corset (n = 30), 2) no corset (n = 30). Workers in the corset group were intended to wear the brace during work for 3 mos. No differences in work-related back injuries or subjective perception of back pain; however, 70% of the corset group felt that the belt aided in avoiding injury, & 29 of 30 said the belt made them "feel good."	No mention of compliance w/ the bracing. Does the brace make the employee feel "overconfident"? No benefit of bracing. However, 97% of corset group said they would continue to wear the brace.
Reddell et al., 1992	II	642 baggage handlers randomized to 4 groups: 1) education, 2) bracing, 3) education & bracing, & 4) no intervention. 58% of group had previous LBI & 26% had a specific LBI. Looked at total incidence of injury, lost or restricted workdays, & Workers' Comp rates over 8-mo period. Found no differences among groups w/ respect to these outcomes. Compliance rate was only 42%.	Authors recommended that back braces not be routinely used for the prevention of LBI in this population. Dropout group had higher incidence of injury than control or education groups (data not included). Even though no benefit was demonstrated, almost 70% of participants found a brace helpful.
Walsh & Schwartz, 1990	II	90 warehouse workers were randomized to 3 groups: 1) no intervention (control), 2) 1 hr of education, or 3) LSO + education. Baseline & 6-mo outcomes assessed as 1) abdominal strength, 2) low-back knowledge assessment, 3) work injury incidence, 4) productivity, & 5) use of health care resources. More than 90% follow-up. No information on brace compliance or participant selection.	Workers in Group 3 missed 2.5 fewer days of work in the 6-mo trial ($p < 0.05$). No differences in productivity. High-risk individuals (those w/ previous back complaints) showed a greater effect, w/ 5.9 fewer lost workdays in Group 3 & 2 fewer in Group 2.
Thompson et al., 1994	III	This was a prospective study of 60 health care transport workers who were divided into 2 groups w/ 41 braced (1 back school intervention) & 19 nonbraced for 3 mos. In workers who were given a back belt, attitudes improved & frequency of back pain decreased.	A 2-part study w/ an initial survey of attitudes in 145 workers who were given a back belt. These workers had better attitudes & a reduction in back injury to 0%. No data are provided for LBP rates & no description is provided for how the authors measured attitude.

* Comp = Compensation; LBI = low-back injury; LBP = low-back pain; LSO = lumbar sacral orthosis; NYC = New York City; pts = patients.

TABLE 2: Bracing for the secondary prevention of low-back pain: summary of evidence*

Authors & Year	Level of Evidence	Description	Comment
Roelofs et al., 2007	I	360 home health care workers w/ a self-reported history of LBP were randomized to 2 groups. The 1st group was allowed to choose 1 of 4 types of braces to be worn during work. The control group received best medical treatment. Participants used a self-reporting calendar to determine the number of days of LBP in a 12-mo period & the number of sick days during the same period. Secondary outcome measures included VAS & QBPDS measures at 3, 6, 9, & 12 mos. Compliance w/ brace was difficult to determine, though the authors state that 78% of participants wore the brace for 1/3 of the days. Overall, 91% returned calendars reporting the number of days of LBP.	Workers in the lumbar support group experienced 52.7 fewer days of LBP (71.7 vs 124.4, $p < 0.001$) but no significant decrease in absentee days (38.5 vs 43.5, $p = 0.45$). There were differences in VAS (4.0 vs 4.6) & functional disability in the previous wk for the lumbar support group. This is a study on the use of bracing for secondary prevention of LBP.
Oleske et al., 2007	II	Randomized clinical trial of auto plant workers who had work-related LBP & were randomized to lumbar support & education (study group) or education alone (control group). Of 868 workers screened, 433 completed at least 1 follow-up visit. Self-reported follow-up was scheduled for 1, 2, 6, & 12 mos. Self-reported outcomes included an LBP & bothersomeness scale (0–10), ODI, & SF-12 physical & mental components; administrative outcomes included medical visits & lost or restricted workdays due to injury or illness. It is uncertain if randomization was attempted on all 868 workers. It is uncertain at what time point the follow-up occurred for the self-reported outcomes for the 433 workers reported on. It is presumed that administrative data are available on all 433.	Both groups improved over 12 mos w/ respect to self-reported outcomes, & there was no difference btwn the groups on any of the self-reported or administrative outcomes. There was a trend to fewer recurrent episodes of LBP in the bracing group (23.1% vs 31.1%), but the greatest reduction was seen in non–assembly line workers.
Jellema et al., 2002	II	Observational study of home health care workers w/ previous LBP. The primary goal was to determine feasibility in a cohort of 62 workers for use of a back brace over 6 mos. Overall, 81% of the participants who had an episode of LBP in the previous wk used the brace. At the end of 6 mos, the VAS score & disability as measured by the QBPDS were both reduced (4.2 vs 2.3 for VAS, 29.3 vs 16.3 for QBPDS).	The authors also used subjective measures of benefit using measures such as satisfaction, confidence that brace reduced pain, & confidence that it provided support. On an NRS, benefit was 7. Although there was benefit, the authors would not recommend brace use in this population.

* LBP = low-back pain; NRS = Numeric Rating Scale; ODI = Oswestry Disability Index; QBPDS = Quebec Back Pain Disability Scale; SF-12 = 12-Item Short Form Health Survey; VAS = visual analog scale.

TABLE 3: Bracing prior to the treatment of low-back pain: summary of evidence*

Authors & Year	Level of Evidence	Description	Comment
Calmels et al., 2009	I	Multicenter randomized trial to evaluate the effect of an elastic lumbar support for subacute LBP. 197 participants were randomized to 1) best medical treatment or 2) lumbar belt + best medical treatment. Primary outcome measures at 30 & 90 days were functional recovery by the EIFEL (RMDQ), change in VAS score, & consumption of analgesic, antiinflammatory, or muscle relaxant medications. The authors state that 90 of 102 pts in the lumbar belt group followed protocol, though there was a trend to wear the brace less as the study continued. Final analysis reported as intention to treat groups.	Pts were recruited from family practitioners. At 30 days, pts in the study group had greater reduction in functional disability (5.6 vs 4.0 on RMDQ, p = 0.02) & VAS (26.8 vs 21.3, p = 0.04) than the control group. These changes continued at 90 days (7.6 vs 6.1, p = 0.02 & 41.5 vs 32.0, p = 0.002). Pharmacologic consumption was reduced as 34.3% of the study group & 56.8% of the control group took medications at 90 days.
Pope et al., 1994	II	164 pts w/ subacute LBP (1–6 mos) randomized to 4 treatment groups: 1) chiropractic manipulation, 2) TMS, 3) massage, & 4) corset. Randomization was 2:1, w/ 70 pts in Group 1 & 31 in the other 3 groups. Pts assessed by VAS, ROM, & lumbar muscle fatigue testing. Full compliance w/ treatment protocol varied btwn groups & was lowest in Group 1 (38%), although 88% of pts completed initial & 3-wk evaluations.	There were no differences in the outcome measures among the groups at 3 wks. All pts were drawn from a chiropractic clinic. No true control group, all pts had an active treatment. Pt confidence was highest in the manipulation group.
Hsieh et al., 1992	II	Pts w/ subacute LBP (3–26 wks) were randomized to manipulation, massage, corset, & TMS. Outcome assessed by ODI & RMDQ. 85 subjects entered into study & 63 (74%) completed questionnaires at the end of 3 wks. The manipulation group performed better than massage & TMS groups, & corset group performed better than massage group.	Authors found both scales reliable for evaluating LBP & response to treatment. This was part of a larger trial. Chiropractic manipulation & corset groups performed better than massage group for both RMDQ & ODI ($p < 0.05$).
Coxhead et al., 1981	III	322 pts randomized to traction, exercises, manipulation, & corset. Pts had sciatic pain w/ or w/o back pain. Factorial study design so that 16 groups were presented in total. Treatment lasted for 4 wks & outcome was assessed at 1, 4, & 16 mos. Outcome measures at 1 mo were improvement on VAS. At all time points, pts given a better, same, or worse satisfaction questionnaire. Also RTW status assessed at 1 & 4 mos. 91% follow-up at 1 mo & 80% at 16 mos.	Pts receiving manipulation had more improvement in VAS at 4 wks ($p < 0.05$). No significant difference in any group at 4 & 16 mos. There is a trend toward subjective improvement in pts receiving more treatments. Active physiotherapy useful in short term.
Million et al., 1981	III	19 pts w/ chronic LBP (>6 mos) were randomized to receive either a soft corset or a corset w/ an insert. Subjective & objective outcomes were compared at 4 & 8 wks. The subjective measure was a 15-item questionnaire that looked at pain & limitation in function as answered by a VAS (Million scale). The objective measurements were ROM & SLR. Summation of objective criteria was used to measure an objective index, though the specific method is not described.	There were no intergroup differences w/ regard to objective criteria, but there was an improvement in pain & function as assessed by the Million scale. Authors concluded that the benefit of a lumbar support does not occur based on an increase in intra-abdominal pressure as evidenced by the lack of improvement in the group w/ the soft binder. The study has a small sample size & short period of use for the brace.

* LBP = low-back pain; ODI = Oswestry Disability Index; pts = patients; RMDQ = Roland-Morris Disability Questionnaire; ROM = range of motion; RTW = return to work; SLR = straight leg raise; TMS = transcutaneous muscle stimulation; VAS = visual analog scale.

TABLE 4: Bracing prior to or following fusion: summary of evidence*

Authors & Year	Level of Evidence	Description	Comment
Yee et al., 2008	II	90 pts undergoing 1-, 2-, or 3-level instrumented posterolateral arthrodesis were randomized to bracing w/ a canvas corset w/ back stays (brace) or no support. Pts were told to stay in brace 24 hrs/day for 8 wks. The primary outcomes were 1- & 2-yr DPQ change. Secondary outcome measures were SF-36, radiographic fusion, & complications. 2 yrs of follow-up in 72 (80%) of 90 cases.	There were no statistically significant btwn-group differences in DPQ or SF-36 at 2 yrs, though both groups had significant improvement over their baseline. No differences were noted for fusion rates or postop complications. Bracing w/ a semi-rigid brace offered no benefits at 1 or 2 yrs over no brace.
Johnsson et al., 1992	III	Comparison of 11 pts w/ rigid external orthosis for 6 mos (Group I) & 11 w/ orthosis for 3 mos (Group II). All pts had a PLF for Grade I or II spondylosis/spondylolisthesis, & tantalum markers were placed at that time. Pts were followed w/ RSA at various time points up to 1 yr. In Group I, 8 of 11 pts had a higher fusion rate based on no translation on RSA compared w/ 2 of 11 pts in Group II.	Movement was assessed by sagittal, vertical, & transverse translation. The motion subsided btwn the 3- & 6-mo exams.
Axelsson et al., 1995	III	All pts who were to undergo fusion for LBP had a trial of rigid or semi-rigid brace therapy for at least 3 wks. Pain improvement was recorded & a significant pain response was judged to be an improvement >50%. Pts w/ a solid radiographic fusion at 1 yr were included in the study (50 pts) & then at 2 yrs judged for pain relief (pain free, significant improvement, slight improvement, unchanged, or worse & as satisfied or as unsatisfied). 31 pts were improved w/ bracing, & 20 of these had a good outcome at 2 yrs while 11 did not. 19 pts did not have significant relief w/ bracing, & 13 of these had favorable outcome at 2 yrs while 6 did not. If bracing is used as a preop test for success after fusion, the sensitivity is 61%, specificity is 35%, PPV is 65%, & NPV is 32%.	The outcome measure used is nonvalidated. The population studied is selected out of a larger pool. The authors compared the percentages of favorable responses w/ & w/o bracing using a chi-square analysis & found no correlation of response to preop bracing & pain relief after solid fusion.

* DPQ = Dallas Pain Questionnaire; LBP = low-back pain; NPV = negative predictive value; PLF = posterolateral lumbar fusion; PPV = positive predictive value; pts = patients; RSA = roentgen stereophotogrammetric analysis; SF-36 = 36-Item Short Form Health Survey.

groups with respect to these outcome measures. Similar to the study by van Poppel and colleagues, only 42% of the individuals in the brace-treated groups were compliant with the use of the brace. The noncompliant group (158 individuals) was followed and found to have a higher incidence of lost workdays following discontinuation of the brace, but the difference between the compliant and noncompliant groups was not significant. This study also provides Level II medical evidence suggesting no benefit for the use of a lumbar orthosis to prevent back injury.

Kraus et al. randomized 12,772 New York City home health attendant workers to 3 groups: 1) lumbar bracing, 2) safety meeting with information, or 3) no intervention at all.[17] The outcome measure was self-reported back injury rates over a period of up to 28 months. The bracing group had fewer episodes of low-back pain than the participants receiving no intervention (rate ratio 1.36, 95% CI 1.02–1.82), and there was a trend toward fewer episodes in the lumbar support group than the information group, although the difference was not significant. Due to randomization techniques and lack of information on the demographic characteristics of the study participants, the follow-up time points reached, and compliance rates, this study offers Level II evidence on the use of bracing as a strategy for primary prevention of low-back pain in home health attendants. The authors also found that the strongest risk factor for low-back injury was a prior back injury, with a 3.1 risk ratio in this population, suggesting that lumbar braces may have an even greater role in secondary prevention of low-back pain.[16,17]

Alexander et al. reported the results of a small prospective randomized study of 60 health care workers divided into 2 groups.[1] One group was assigned to wear a lumbar corset for a 3-month period. No differences in work-related back injuries or perception of back pain were noted. This study was downgraded to Level II evidence due to the use of a nonvalidated outcome measure but does suggest that a corset-type orthosis is not an effective measure to prevent low-back pain.

Walsh and Schwartz reported on a group of 90 warehouse workers who were randomly assigned to 3 groups: 1) no intervention; 2) 1-hour education; or 3) 6-month lumbosacral molded semi-rigid orthosis therapy and education.[36] Outcomes were assessed using various measures, including work injury incidence, work productivity, and utilization of health care resources. Brace-treated workers missed 2.5 days less work ($p = 0.03$) than those not wearing braces (control and education-only groups), but there were no statistically significant differences between the groups with respect to productivity or utilization of health care resources. A subgroup analysis revealed that the benefit in terms of number of lost workdays was greatest in patients with a previous back injury. The authors concluded that the combination of brace therapy and education was effective in reducing lost workdays, especially among patients with a history of back injury. Limitations of this study include failure to incorporate validated outcome measures and failure to describe worker compliance with the bracing routine. Therefore, this study is considered to provide Level II evidence in support of brace therapy as an alternative for prevention of low-back pain.

Post hoc analysis from many of the initial studies on the efficacy of bracing for the prevention of low-back pain (primary prevention) revealed that the strongest benefit for lumbar bracing was derived from workers with a prior history of low-back pain.[1,16,17,36] Therefore, more recent studies have been designed to look specifically at the utility of lumbar bracing in workers with a prior history of low-back pain (secondary prevention of low-back pain). Roelofs et al. studied the use of bracing in 360 home health workers with a history of back pain, defined as current back pain or 2 or more episodes of low-back pain in the previous year.[29] Workers were assigned to a short course on healthy working methods with or without use of a brace. Over 12 months, the group of workers who were assigned to use of a brace had 52.7 fewer self-reported days with low-back pain (95% CI –59.6 to –45.1), but there was no statistically significant difference between the groups in days of sick leave (38.5 vs 43.5, 95% CI –21.1 to 6.8). Secondary outcome measures included VAS, Quebec Back Pain Disability Scale measures, and self-reported low-back pain–related sick days at 3, 6, 9, and 12 months. The bracing group had a lower mean VAS for low-back pain (4.0 vs 4.6, $p = 0.02$), better mean disability rating (26.2 vs 30.3, $p = 0.017$), and fewer days of low-back pain–related sick leave (3.2 vs 8.0, $p = 0.003$). The use of a back brace was at the discretion of the worker and only a rough estimate of use was given, suggesting the workers used the brace about one-third of the time, although the authors report an adherence rate of 78%. The baseline characteristics of the 2 groups were very similar, and 91% of participants returned self-reported low-back pain calendars. Therefore, this study is considered to provide Level I evidence on the benefits of the prescription of bracing for limiting the number of days of low-back pain in home health workers with a prior history of low-back pain.

Oleske and colleagues performed a randomized clinical trial involving auto plant workers who had work-related low-back pain and were randomly assigned to lumbar support and education (study group) or education alone (control group).[25] Of 868 workers screened, 433 workers completed at least 1 follow-up visit. Self-reported outcome follow-up was scheduled for 1, 2, 6, and 12 months. Self-reported outcome measures included a low-back pain and bothersomeness scale (0–10), the Oswestry Disability Index (ODI), and the physical and mental components of the 12-Item Short Form Health Survey (SF-12); administrative outcomes included medical visits and lost or restricted workdays due to injury or illness. It is uncertain whether randomization was attempted for all 868 workers. With respect to the 433 participants on whom the authors reported, it is uncertain at what time point the follow-up occurred for the self-reported outcomes. It is presumed that administrative data are available for all 433 participants. Both groups reported significant declines in low-back pain (VAS), disability (ODI), and neurogenic symptoms and improvement in overall physical health (SF-12 scores) over 12 months. There was no significant difference in the number of lost or restricted workdays between the groups. There was a trend toward fewer episodes of low-back pain in the brace group (23.1%

vs 31.1%, p = 0.059). A subgroup analysis showed a significant decline in the number of recurrent episodes in the non–assembly line workers receiving a brace (34.9% vs 63.1%, p = 0.016). Because of the uncertainties in the randomization, the dropout rate of 50%, and the lack of clarity regarding the number of workers who achieved 6 or 12 months of follow-up, this study is considered to provide Level II evidence that braces have no impact on lost work time, disability, or medical utilization in a general working population.

Jellema et al. performed an observational study on a cohort of home health care workers who had previous low-back pain.[12] The primary goal was to determine feasibility in a cohort of 62 workers for use of a back brace over 6 months. Overall, 81% of the participants who had an episode of low-back pain in the previous week used the brace. At the end of 6 months, the authors observed a 44% reduction in both the mean VAS pain score (4.2 vs 2.3) and the mean disability score as measured by the Quebec Back Pain Disability Scale (29.3 vs 16.3). Although there was a dropout rate of 20% due to a relatively small sample size, the study provides Level II evidence that bracing is a feasible option in home health care workers with prior low-back pain. The authors, however, recommended a prospective randomized trial to further determine the role of bracing in this population.

Several historical cohort studies have examined the incidence of back pain and days lost to work in groups of workers before and after they were issued a brace or lumbar support belt by their employer. Analysis of these studies revealed mixed results. One study identified no change in the incidence of back pain and sick days after braces were issued, and 2 studies reported a reduction in these parameters following the issue of a lumbar support to employees.[16,23,30] Overall, the medical evidence supporting the use of braces for prevention of low-back pain is inconsistent. The authors of several systematic literature reviews have concluded that lumbar support devices are not useful for the prevention of low-back pain in the general working population.[13,32,35] It does appear, however, that braces may be useful as a measure to decrease the number of sick days lost due to low-back pain in workers with a history of low-back injury (secondary prevention).

Bracing for the Treatment of Low-Back Pain

There have been several randomized control trials investigating the role of bracing as a treatment for low-back pain. A multicenter randomized trial by Calmels et al. evaluated the effect of an elastic lumbar support for subacute low-back pain.[5] One hundred ninety-seven participants were randomized to best medical treatment or best medical treatment supplemented with the elastic lumbar support. Primary outcome measures at 30 and 90 days were functional recovery by the EIFEL (French version of the Roland-Morris Disability Questionnaire [RMDQ]), change in pain VAS score, and consumption of analgesic and anti-inflammatory medications or muscle relaxants. At 30 days, patients in the study group had greater reduction in functional disability (5.6 vs 4.0 on RMDQ, p = 0.02) and VAS (26.8 vs. 21.3, p = 0.04) than the control group. These changes continued at 90 days (7.6 vs 6.1, p = 0.02, and 41.5 vs 32.0, p = 0.002). Consumption of pharmaceutical agents was reduced, as 34.3% of the study group and 56.8% of the control group took medication at 90 days. There were few limitations identified within the study design and execution, and therefore this study is considered to provide Level I medical evidence in support of bracing for the short-term management of subacute low-back pain.

Valle-Jones and colleagues randomized 216 patients with nonspecific low-back pain of varying duration to lumbar brace therapy or activity modification for 3 weeks.[31] Outcome measures included a VAS score for pain and disability. Patients were also asked to record usage of pain medication. Brace-treated patients were found to have more improvement in pain at rest, pain with activity, and pain at night between Days 7 and 21. In addition, brace-treated patients took half the number of doses of paracetamol during the 21-day trial period compared with the control group. Return-to-work rates were higher in the brace-treated group (85%) than in the control group (67%, p < 0.02). The inclusion of diverse patient populations (acute and chronic low-back pain), the use of nonvalidated outcome measures (a 7-point VAS), and lack of data detract from the trial. This paper is considered to provide Level II medical evidence supporting the efficacy of braces for the short-term amelioration of low-back pain.

Pope et al. studied 164 patients with low-back pain drawn from a chiropractic clinic. Patients were randomized to 4 treatments: 1) chiropractic manipulation; 2) transcutaneous muscle stimulation (TMS); 3) massage; and 4) lumbar corset.[26] Patients were assessed for pain using a VAS and were also assessed for range of motion after 3 weeks of treatment. There were no differences among the groups. Because of the relatively small treatment groups (~ 30 patients in 3 of the 4 groups) and selected patient population (from a chiropractic practice), this paper is considered to provide Level II medical evidence suggesting that braces are no more effective than other modalities used for the treatment of acute low-back pain. Hsieh et al. studied 63 patients with low-back pain of less than 6 months' duration.[11] Patients were randomized to manipulation, massage, lumbar corset, or TMS treatment for 3 weeks. Functional outcomes were assessed with the ODI and RMDQ. The primary purpose of the study was to validate the disability scales. Chiropractic manipulation and corset performed better than massage for both RMDQ and ODI (p < 0.05). The small number of patients in each cohort and the lack of a power analysis limit the authors' conclusions. This paper provides Level II evidence supporting the role of short-term lumbar brace therapy in patients with acute or subacute low-back pain as compared with massage or TMS. No inferences can be drawn regarding the effect of braces for patients with chronic low-back pain.

Two randomized controlled studies published in 1981 provide information on lumbar brace therapy for low-back pain. Coxhead and coworkers performed a randomized study of 322 sciatica patients with or without low-back pain randomized to different treatment modalities, including traction, exercises, manipulation, corset brace, and combinations of these treatments for a total

of 16 treatment groups.[9] Treatments lasted for 4 weeks, and outcome was assessed at 1, 4, and 16 months by VAS, return-to-work status, and patient satisfaction criteria. No benefit, short or long term, was detected for the use of lumbar corset braces. Because the population was composed of patients with sciatica, no direct conclusions can be drawn with regard to the treatment of low-back pain. In a smaller cohort study of 19 patients with chronic low-back pain, Million et al. randomized patients to either a soft or rigid lumbar brace group for 4 weeks.[21] A 15-item questionnaire about pain and functional limitation on a VAS (Million scale) demonstrated a significant improvement ($p = 0.01$) for the cohort of patients wearing a rigid brace at 4 and 8 weeks. Rigid lumbar bracing may therefore have some short-term benefit compared with soft bracing for the short-term treatment of low-back pain. Because there was no control group in this study, the paper is considered to provide Level III medical evidence regarding the efficacy of brace therapy for low-back pain.

Bracing Prior to Fusion

There has only been one study published that has investigated the role of preoperative brace therapy as a predictor for outcome following lumbar fusion.[4] Axelsson et al. placed all patients who were scheduled to undergo a lumbar fusion for low-back pain in either a rigid or a semi-rigid brace for at least 3 weeks. Pain improvement was recorded, and 31 patients had a significant response, judged as an improvement in pain of at least 50%. Only 50 patients with a solid radiographic posterolateral fusion on anteroposterior and lateral plain radiographs at 1 year were included in the study. Two years following surgery, these same patients were subjectively examined for pain relief and satisfaction. Of the 31 patients who had experienced significant improvement of pain with the preoperative corset, 20 had a good outcome at 2 years (pain free or significant improvement), whereas 11 patients had poor outcomes despite a favorable response to preoperative lumbar bracing. Nineteen patients did not have significant relief from the corset, and 13 of these reported a favorable outcome at 2 years. If lumbar bracing is used as a preoperative "prognostic test" for success after solid fusion, the sensitivity is 61%, the specificity is 35%, the positive predictive value is 65%, and the negative predictive value is 32%. Therefore, due to the poor diagnostic parameters, the use of lumbar bracing as a prognostic indicator of fusion outcome is not recommended. Because of the reliance on patient satisfaction scores, the select population studied (only patients with solid radiographic fusion), and the lack of a standardized bracing protocol, the medical evidence derived from this study is considered Level III.

Bracing Following Fusion

Until recently, there were no published studies that compared outcomes following lumbar fusion with and without the supplemental use of a lumbar orthosis. Yee et al. randomized 90 patients undergoing 1-, 2-, or 3-level instrumented PLF to 8 weeks of postoperative bracing with a canvas corset with back stays (brace) or no orthosis.[37] Data from 1- and 2-year follow-up examinations were available for 72 (80%) of the 90 patients. There were no statistically significant between-group differences in Dallas Pain Questionnaire (DPQ) or SF-36 results at 1 or 2 years, although both groups showed significant improvement compared with baseline. No differences were noted for fusion rates or postoperative complications. Due to the good compliance and follow-up rates and an appropriate study size based on the power calculation, this study is considered to provide Level I evidence that postoperative semi-rigid bracing offers no functional or radiographic benefit at 1 or 2 years after surgery for patients undergoing instrumented PLF.

Several authors have advocated the use of brace therapy following lumbar fusion surgery.[8,14] Johnsson et al. have suggested a minimum 5-month period of bracing following noninstrumented lumbar fusion.[15] They noted that patients who used a brace for 6 months following surgery had a higher fusion rate (8 of 11 patients) at 1 year than those who used a brace for 3 months (2 of 11), when fusion was assessed as lack of motion with roentgen stereophotogrammetric analysis (RSA). The authors found that sagittal and vertical translation decreased significantly as measured by RSA between 3 and 6 months following surgery. They interpreted this result as evidence that healing of a noninstrumented lumbar fusion occurs over a 6-month period. They presented no evidence, however, regarding the effect of lumbar bracing on the rate of lumbar spinal fusion or functional outcome.

Summary

Although conflicting reports have been presented in the literature regarding the utility of lumbar bracing for the prevention of low-back pain, lower-level evidence suggests that the prophylactic use of braces does not reduce the incidence of low-back pain or decrease the amount of lost productivity in the general working population. In the select population of workers with a history of a back injury, bracing appears to decrease the number of workdays lost due to back pain.

Lumbar bracing appears to be an effective treatment for acute low-back pain in select populations. They do not appear to be an effective treatment strategy for chronic low-back pain. If a brace is used, rigid braces offer some benefit over soft braces. There are no data to suggest that relief of low-back pain with preoperative external bracing predicts a favorable outcome following lumbar spinal fusion. Bracing following instrumented lumbar fusion for degenerative disease does not appear to improve fusion rates or clinical outcomes.

Key Issues for Future Investigation

The most relevant questions for the spine surgeon may be related to the predictive value of a trial of brace therapy to predict functional outcomes following lumbar fusion surgery and the ability of postoperative bracing to improve functional and radiographic outcomes of fusion surgery. Formalizing and performing an appropriate prognostic study to investigate the predictive value of bracing may prove to be too difficult to perform. To determine

the efficacy of postoperative bracing, an RCT comparing patients undergoing similar lumbar fusion procedures, randomized to brace therapy or no such therapy, could provide additional high-quality evidence to address the effect of postoperative bracing on functional and radiographic outcome, although the sample size would have to be large to demonstrate a small improvement in outcome.

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., 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 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: Dailey. 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: Dailey. Study supervision: Kaiser.

References

1. Alexander A, Woolley SM, Bisesi M, Schaub E: The effectiveness of back belts on occupational back injuries and worker perception. **Professional Safety 40:**22–26, 1995
2. Axelsson P, Johnsson R, Strömqvist B: Effect of lumbar orthosis on intervertebral mobility. A roentgen stereophotogrammetric analysis. **Spine (Phila Pa 1976) 17:**678–681, 1992
3. Axelsson P, Johnsson R, Strömqvist B: Lumbar orthosis with unilateral hip immobilization. Effect on intervertebral mobility determined by roentgen stereophotogrammetric analysis. **Spine (Phila Pa 1976) 18:**876–879, 1993
4. Axelsson P, Johnsson R, Strömqvist B, Nilsson LT, Akesson M: Orthosis as prognostic instrument in lumbar fusion: no predictive value in 50 cases followed prospectively. **J Spinal Disord 8:**284–288, 1995
5. Calmels P, Queneau P, Hamonet C, Le Pen C, Maurel F, Lerouvreur C, et al: Effectiveness of a lumbar belt in subacute low back pain: an open, multicentric, and randomized clinical study. **Spine (Phila Pa 1976) 34:**215–220, 2009
6. Cholewicki J: The effects of lumbosacral orthoses on spine stability: what changes in EMG can be expected? **J Orthop Res 22:**1150–1155, 2004
7. Cholewicki J, Shah KR, McGill KC: The effects of a 3-week use of lumbosacral orthoses on proprioception in the lumbar spine. **J Orthop Sports Phys Ther 36:**225–231, 2006
8. Connolly PJ, Grob D: Bracing of patients after fusion for degenerative problems of the lumbar spine—yes or no? **Spine (Phila Pa 1976) 23:**1426–1428, 1998
9. Coxhead CE, Inskip H, Meade TW, North WR, Troup JD: Multicentre trial of physiotherapy in the management of sciatic symptoms. **Lancet 1:**1065–1068, 1981
10. Dillingham TR: Lumbar supports for prevention of low back pain in the workplace. **JAMA 279:**1826–1828, 1998
11. Hsieh CY, Phillips RB, Adams AH, Pope MH: Functional outcomes of low back pain: comparison of four treatment groups in a randomized controlled trial. **J Manipulative Physiol Ther 15:**4–9, 1992
12. Jellema P, Bierma-Zeinstra SM, Van Poppel MN, Bernsen RM, Koes BW: Feasibility of lumbar supports for home care workers with low back pain. **Occup Med (Lond) 52:**317–323, 2002
13. Jellema P, van Tulder MW, van Poppel MN, Nachemson AL, Bouter LM: Lumbar supports for prevention and treatment of low back pain: a systematic review within the framework of the Cochrane Back Review Group. **Spine (Phila Pa 1976) 26:** 377–386, 2001
14. Johnsson R: The use of orthoses in lumbar spine fusion. **Acta Orthop Scand Suppl 251:**92–93, 1993
15. Johnsson R, Strömqvist B, Axelsson P, Selvik G: Influence of spinal immobilization on consolidation of posterolateral lumbosacral fusion. A roentgen stereophotogrammetric and radiographic analysis. **Spine (Phila Pa 1976) 17:**16–21, 1992
16. Kraus JF, Brown KA, McArthur DL, Peek-Asa C, Samaniego L, Kraus C: Reduction of acute low back injuries by use of back supports. **Int J Occup Environ Health 2:**264–273, 1996
17. Kraus JF, Schaffer KB, Rice T, Maroosis J, Harper J: A field trial of back belts to reduce the incidence of acute low back injuries in New York City home attendants. **Int J Occup Environ Health 8:**97–104, 2002
18. Lahad A, Malter AD, Berg AO, Deyo RA: The effectiveness of four interventions for the prevention of low back pain. **JAMA 272:**1286–1291, 1994
19. Lantz SA, Schultz AB: Lumbar spine orthosis wearing. I. Restriction of gross body motions. **Spine (Phila Pa 1976) 11:** 834–837, 1986
20. Lantz SA, Schultz AB: Lumbar spine orthosis wearing. II. Effect on trunk muscle myoelectric activity. **Spine (Phila Pa 1976) 11:**838–842, 1986
21. Million R, Nilsen KH, Jayson MI, Baker RD: Evaluation of low back pain and assessment of lumbar corsets with and without back supports. **Ann Rheum Dis 40:**449–454, 1981
22. Minor SD: Use of back belts in occupational settings. **Phys Ther 76:**403–408, 1996
23. Mitchell LV, Lawler FH, Bowen D, Mote W, Asundi P, Purswell J: Effectiveness and cost-effectiveness of employer-issued back belts in areas of high risk for back injury. **J Occup Med 36:**90–94, 1994
24. Nachemson AL: Orthotic treatment for injuries and diseases of the spinal column. **Phys Med Rehabil 1:**11–24, 1987
25. Oleske DM, Lavender SA, Andersson GB, Kwasny MM: Are back supports plus education more effective than education alone in promoting recovery from low back pain?: Results from a randomized clinical trial. **Spine (Phila Pa 1976) 32:**2050–2057, 2007

26. Pope MH, Phillips RB, Haugh LD, Hsieh CY, MacDonald L, Haldeman S: A prospective randomized three-week trial of spinal manipulation, transcutaneous muscle stimulation, massage and corset in the treatment of subacute low back pain. **Spine (Phila Pa 1976) 19:**2571–2577, 1994
27. Reddell CR, Congleton JJ, Huchingson RD, Montgomery JF: An evaluation of a weightlifting belt and back injury prevention training class for airline baggage handlers. **Appl Ergon 23:**319–329, 1992
28. 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 14: brace therapy as an adjunct to or substitute for lumbar fusion. **J Neurosurg Spine 2:**716–724, 2005
29. Roelofs PD, Bierma-Zeinstra SM, van Poppel MN, Jellema P, Willemsen SP, van Tulder MW, et al: Lumbar supports to prevent recurrent low back pain among home care workers: a randomized trial. **Ann Intern Med 147:**685–692, 2007
30. Thompson L, Pati AB, Davidson H, Hirsh D: Attitudes and back belts in the workplace. **Work 4:**22–27, 1994
31. Valle-Jones JC, Walsh H, O'Hara J, O'Hara H, Davey NB, Hopkin-Richards H: Controlled trial of a back support ('Lumbotrain') in patients with non-specific low back pain. **Curr Med Res Opin 12:**604–613, 1992
32. van Duijvenbode IC, Jellema P, van Poppel MN, van Tulder MW: Lumbar supports for prevention and treatment of low back pain. **Cochrane Database Syst Rev (2):**CD001823, 2008
33. van Poppel MN, de Looze MP, Koes BW, Smid T, Bouter LM: Mechanisms of action of lumbar supports: a systematic review. **Spine (Phila Pa 1976) 25:**2103–2113, 2000
34. van Poppel MN, Koes BW, Devillé W, Smid T, Bouter LM: Risk factors for back pain incidence in industry: a prospective study. **Pain 77:**81–86, 1998
35. Van Tulder MW, Jellema P, van Poppel MN, Nachemson AL, Bouter LM: Lumbar supports for prevention and treatment of low back pain. **Cochrane Database Syst Rev (3):**CD001823, 2000
36. Walsh NE, Schwartz RK: The influence of prophylactic orthoses on abdominal strength and low back injury in the workplace. **Am J Phys Med Rehabil 69:**245–250, 1990
37. Yee AJ, Yoo JU, Marsolais EB, Carlson G, Poe-Kochert C, Bohlman HH, et al: Use of a postoperative lumbar corset after lumbar spinal arthrodesis for degenerative conditions of the spine. A prospective randomized trial. **J Bone Joint Surg Am 90:**2062–2068, 2008

Manuscript submitted March 17, 2014.
Accepted April 9, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14282.

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.

J Neurosurg Spine 21:102–105, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 15: Electrophysiological monitoring and lumbar fusion

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

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

Intraoperative monitoring (IOM) is commonly used during lumbar fusion surgery for the prevention of nerve root injury. Justification for its use stems from the belief that IOM can prevent nerve root injury during the placement of pedicle screws. A thorough literature review was conducted to determine if the use of IOM could prevent nerve root injury during the placement of instrumentation in lumbar or lumbosacral fusion. There is no evidence to date that IOM can prevent injury to the nerve roots. There is limited evidence that a threshold below 5 mA from direct stimulation of the screw can indicate a medial pedicle breach by the screw. Unfortunately, once a nerve root injury has taken place, changing the direction of the screw does not alter the outcome. The recommendations formulated in the original guideline effort are neither supported nor refuted with the evidence obtained with the current studies. *(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14324)*

Key Words • **intraoperative monitoring** • **fusion** • **lumbar spine** • **practice guidelines**

Recommendations

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

Grade I

The use of direct screw stimulation evoked electromyography (EMG) responses, as a diagnostic modality during lumbar fusion surgery, is an option since evidence suggests that EMG monitoring can be highly sensitive in detecting breaches of the pedicle (one Level III study).

The data are insufficient to support a recommendation regarding the use of neuromonitoring as a modality that can be used for the preservation of nerve root function during lumbar fusion surgery (one Level IV study).

Rationale

Intraoperative monitoring (IOM) is commonly used during spinal deformity surgery and resection of intramedullary tumors, as well as other nonspine surgeries including repair of aortic aneurysms.[2–8] The use of IOM during routine surgery for degenerative lumbar disease remains controversial; however, supporters of IOM claim that this modality enhances the placement of pedicle screws. Based on the review from the original guidelines, there is relatively good evidence that the use of IOM provides useful information pertaining to the integrity of the pedicle wall and the potential for neurological injury during pedicle screw insertion.[11]

Abbreviations used in this paper: EMG = electromyography; IOM = intraoperative monitoring; MEP = motor evoked potential; SSEP = somatosensory evoked potential.

Several important questions pertaining to the use of IOM during lumbar fusion surgery remain unanswered and include the following:

1) Does intraoperative electrophysiological monitoring of the nerve roots increase the safety of lumbar or lumbosacral instrumentation?

2) Does the use of intraoperative electrophysiological monitoring influence patient outcomes following lumbar spine fusion surgery for degenerative disease?

The current literature review was intended to address these queries and examine the evidence pertaining to the utility of IOM during lumbar fusion surgery for degenerative disease.

Search Criteria

A computerized search of the database of the National Library of Medicine from 2004 to December 2011 was conducted using the search terms ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ("Electrophysiology"[MeSH] OR "Evoked Potentials"[MeSH] OR "electromyography"[MeSH]). The search was restricted to the English language and human subjects, yielding a total of 89 citations. The titles and abstracts of each of these references were reviewed, and papers not concerned with the use of monitoring for lumbosacral fusion were removed. The references that provided either direct or supporting evidence relevant to the use of monitoring for lumbar or lumbosacral fusion procedures were included for review. Relevant references from the bibliographies of these papers were also identified and listed. Since the previous guidelines publication, 3 new articles have been published that specifically address the role of IOM in lumbar fusion. Two studies examined the role of neuromonitoring in thoracolumbar procedures as well as decompressive procedures.[1,10] One published case report reported injury to the iliac artery that was detected by IOM.[9]

Scientific Foundation

Under ideal circumstances, the use of IOM would allow the surgeon to perform the intended procedure with less risk and provide information predictive of outcome. Since the publication of the original guidelines, there have been relatively few studies published that provide further insight into the utility of IOM for procedures to treat degenerative disease of the lumbar spine. The recommendations published in the original guidelines support the use of IOM, both somatosensory evoked potential (SSEP) and EMG, when the surgeon desires immediate intraoperative feedback regarding the potential of neurological injury and/or immediate feedback regarding the integrity of the pedicle wall when internal stabilization is intended with pedicle screws.[11]

Alemo and Sayadipour[1] performed a retrospective study in 86 patients who underwent lumbar fusion (37 patients) or lumbosacral fusion (49 patients), all with the placement of titanium pedicle screws (Table 1). Somatosensory evoked potential, motor evoked potential (MEP), and evoked EMG testing of pedicle screws were performed. In their study, 28 (5%) of 414 screws were found to have a response with evoked EMG testing intraoperatively. All of these screws were repositioned, and none of these patients were found to have a postoperative neurological deficit. There were 3 false-negative EMG evoked responses during surgery. These were discovered after the patients woke up with a new neurological deficit. Unfortunately, the misplacement of the screws was detected by postoperative CT scanning and not through neuromonitoring. Based on this study there is no evidence to suggest that intraoperative neuromonitoring can be used to prevent neurological deficits during surgery.

Parker et al.[10] performed a retrospective study examining the records of 418 patients in whom 2450 consecutive pedicle screws were placed (Table 1). Multimodality neuromonitoring was performed (MEPs, SSEPs, and evoked EMG response) for all surgeries that were performed on the lumbar spine (L1–S1). This study was unique in that CT scans were obtained 48 hours after the surgery to confirm placement of the screws. Screw positions on CT scans were correlated to EMG evoked responses during surgery. A response below 7 mA indicated to the surgeon that there might be malpositioning of the screw. It is unclear from the paper the number of screws that were repositioned during surgery. Overall there was a 0.7% false-negative rate (intraoperatively the screw demonstrated no stimulation below 10 mA while it was found to have a medial breach on CT scanning). The authors correlated the EMG evoked responses to the position of the screw on the CT scan to determine if there was a particular threshold. In this study, the authors were able to demonstrate that an EMG evoked response below 5 mA had a low sensitivity (43.4%) but high specificity (99.9%) in detecting a medial breach of the pedicle screw. This study supports previous literature that supports the use of EMG testing during placement of instrumentation in lumbar fusion procedures. Unfortunately, the paper could not demonstrate any neuromonitoring findings that could be used to help the surgeon avoid neurological injury during placement of the instrumentation.

The Use of Neuromonitoring During Anterior Lumbar Fusion

The majority of publications investigating the utility of IOM with anterior lumbar procedures have been case reports, limiting the strength of the data and any conclusions that may be formulated. In one published case report, there was a loss of MEP and SSEP signal to the left lower extremity during surgery that correlated to occlusion of the left iliac artery.[9] Intraoperative exploration revealed that the iliac artery had become trapped within the L4–5 disc space. Following a release of the artery, a full recovery of signal was observed and no neurological deficits were observed following the procedure. Although this evidence is purely anecdotal, at best Level IV evidence, this study provides an example of the use of IOM identified a potential injury that was correctable.

TABLE 1: Electrophysiological monitoring and lumbar fusion: summary of evidence

Authors & Year	Description	Results	Level of Evidence	Conclusions
Alemo & Sayadipour, 2010	Retrospective study that examined the use of pedicle screws in lumbar fusion.	28/414 screws (5%) evoked an EMG response during surgery. All of these screws were repositioned & there were no permanent neurological deficits associated w/ these screws. Three false-negative evoked EMGs (3.48%) were detected. These patients woke up w/ a new neurological deficit.	IV, retrospective case series	The use of intraop neuromonitoring during lumbar fusion surgery cannot be routinely recommended as it does not always detect new postop neurological deficits.
Parker et al., 2011	Retrospective review of the placement of 2450 pedicle screws. Intraop neuromonitoring was correlated to postop CT scans.	An EMG evoked response <5 mA demonstrated a low specificity & high sensitivity in detecting a medial breach of the pedicle screw.	III, retrospective case series	This paper does not provide the surgeon any information that can be used during surgery to avoid neurological deficits during the placement of lumbar pedicle screws.

Summary

The current literature review provided no new high-quality studies supporting the use of IOM during lumbar fusion for degenerative spine disease. The routine use of IOM for this type of surgery, therefore, cannot be recommended. The recommendations formulated in the original guideline effort are neither supported nor refuted with the evidence obtained with the current studies.

Several low-quality studies demonstrated a correlation between changes in SSEP signals and nerve root injury. Unfortunately, once a change has occurred, there is no evidence to suggest that intraoperative maneuvers can lead to recovery of the nerve function. There is evidence to suggest that a threshold below 5 mA indicates a medial breach of the pedicle screw, although it is unclear how this affects the overall outcome. Finally, there is no evidence to suggest that neurophysiological monitoring during lumbar spine fusion can alter the outcome of surgery. Unfortunately, the recent literature does little to address the concerns previously stated.

Key Issues for Future Investigation

To date, there has been no randomized, prospective, multicenter trial that has examined the value of IOM during lumbar fusion surgery. Investigating the utility of IOM may prove impractical, as the true value of intraoperative signal changes could only be determined through a study in which a cohort of patients received no intervention for alternations in IOM observed during surgery. Such a study would in all likelihood be considered unethical. Such information, however, will be essential to perform a validated cost-effectiveness analysis to determine whether the benefits of IOM justify the added cost.

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., 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 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: Sharan. 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: Sharan. Study supervision: Kaiser.

References

1. Alemo S, Sayadipour A: Role of intraoperative neurophysiologic monitoring in lumbosacral spine fusion and instrumentation: a retrospective study. **World Neurosurg 73:**72–76, 2010
2. Brown RH, Nash CL Jr, Berilla JA, Amaddio MD: Cortical evoked potential monitoring. A system for intraoperative

monitoring of spinal cord function. **Spine (Phila Pa 1976) 9:**256–261, 1984
3. Dawson EG, Sherman JE, Kanim LE, Nuwer MR: Spinal cord monitoring. Results of the Scoliosis Research Society and the European Spinal Deformity Society survey. **Spine (Phila Pa 1976) 16 (8 Suppl):**S361–S364, 1991
4. Dinner DS, Lüders H, Lesser RP, Morris HH, Barnett G, Klem G: Intraoperative spinal somatosensory evoked potential monitoring. **J Neurosurg 65:**807–814, 1986
5. Jellinek D, Jewkes D, Symon L: Noninvasive intraoperative monitoring of motor evoked potentials under propofol anesthesia: effects of spinal surgery on the amplitude and latency of motor evoked potentials. **Neurosurgery 29:**551–557, 1991
6. Keith RW, Stambough JL, Awender SH: Somatosensory cortical evoked potentials: a review of 100 cases of intraoperative spinal surgery monitoring. **J Spinal Disord 3:**220–226, 1990
7. Loder RT, Thomson GJ, LaMont RL: Spinal cord monitoring in patients with nonidiopathic spinal deformities using somatosensory evoked potentials. **Spine (Phila Pa 1976) 16:** 1359–1364, 1991
8. Macon JB, Poletti CE, Sweet WH, Ojemann RG, Zervas NT: Conducted somatosensory evoked potentials during spinal surgery. Part 2: clinical applications. **J Neurosurg 57:**354–359, 1982
9. Nair MN, Ramakrishna R, Slimp J, Kinney G, Chesnut RM: Left iliac artery injury during anterior lumbar spine surgery diagnosed by intraoperative neurophysiological monitoring. **Eur Spine J 19 (Suppl 2):**S203–S205, 2010
10. Parker SL, Amin AG, Farber SH, McGirt MJ, Sciubba DM, Wolinsky JP, et al: Ability of electromyographic monitoring to determine the presence of malpositioned pedicle screws in the lumbosacral spine: analysis of 2450 consecutively placed screws. Clinical article. **J Neurosurg Spine 15:**130–135, 2011
11. 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 15: electrophysiological monitoring and lumbar fusion. **J Neurosurg Spine 2:**725–732, 2005

Manuscript submitted March 27, 2014.
Accepted April 9, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14324.
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.

J Neurosurg Spine 21:106–132, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 16: Bone graft extenders and substitutes as an adjunct for lumbar fusion

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

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

In an attempt to enhance the potential to achieve a solid arthrodesis and avoid the morbidity of harvesting autologous iliac crest bone (AICB) for a lumbar fusion, numerous alternatives have been investigated. The use of these fusion adjuncts has become routine despite a lack of convincing evidence demonstrating a benefit to justify added costs or potential harm. Potential alternatives to AICB include locally harvested autograft, calcium-phosphate salts, demineralized bone matrix (DBM), and the family of bone morphogenetic proteins (BMPs). In particular, no option has created greater controversy than the BMPs. A significant increase in the number of publications, particularly with respect to the BMPs, has taken place since the release of the original guidelines. Both DBM and the calcium-phosphate salts have demonstrated efficacy as a graft extender or as a substitute for AICB when combined with local autograft. The use of recombinant human BMP-2 (rhBMP-2) as a substitute for AICB, when performing an interbody lumbar fusion, is considered an option since similar outcomes have been observed; however, the potential for heterotopic bone formation is a concern. The use of rhBMP-2, when combined with calcium phosphates, as a substitute for AICB, or as an extender, when used with local autograft or AICB, is also considered an option as similar fusion rates and clinical outcomes have been observed. Surgeons electing to use BMPs should be aware of a growing body of literature demonstrating unique complications associated with the use of BMPs.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14325)

Key Words • **lumbar spine** • **bone graft** • **bone substitute** • **fusion** • **bone morphogenetic protein** • **practice guidelines**

Abbreviations used in this paper: ACS = absorbable collagen sponge; AICB = autologous iliac crest bone; ALIF = anterior lumbar interbody fusion; β-TCP = β-tricalcium phosphate; BMA = bone marrow aspirate; CHA = coralline hydroxyapatite; CRM = compression-resistant matrix; DBM = demineralized bone matrix; FRA = femoral ring allograft; HA = hydroxyapatite; ICBG = iliac crest bone graft; IDE = investigational device exemption; mJOA = modified Japanese Orthopaedic Association; NRS = numeric rating scale; ODI = Oswestry Disability Index; OP-1 = osteogenic protein–1; PLIF = posterior lumbar interbody fusion; RCT = randomized controlled trial; rhBMP = recombinant human bone morphogenetic protein; SF-36 = 36-Item Short Form Health Survey; TLIF = transforaminal lumbar interbody fusion; VAS = visual analog scale.

Recommendations

There is no evidence that conflicts with the previous recommendations published in the original version of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine" regarding the use of hydroxyapatite (HA), various calcium-based preparations, and recombinant human bone morphogenetic protein–2 (rhBMP-2) as bone graft extenders and substitutes for lumbar fusion.[48]

No prior recommendations regarding the use of rhBMP-7 for lumbar fusions were published in the original Lumbar Fusion Guidelines.

Demineralized Bone Matrix

Grade C (Single Level III and Single Level V Studies)

The use of demineralized bond matrix (DBM) as a bone graft extender is an option for 1- and 2-level instrumented posterolateral fusions.

Hydroxyapatite/Calcium Extenders

Grade C (Single Level II Study)

The use of β-tricalcium phosphate (β-TCP)/local autograft as a substitute for autologous iliac crest bone (AICB) is an option for single-level instrumented posterolateral fusion due to comparable fusion rates and clinical outcomes.

Grade C (Single Level II Study)

The use of HA with local autograft/bone marrow aspirate (BMA) as a substitute for AICB in an option for instrumented posterolateral fusion due to comparable fusion rates and clinical outcomes.

Grade C (Multiple Level V Studies)

The use of HA can be considered an option as a graft extender when mixed with AICB for instrumented posterolateral fusions.

Grade C (Single Level IV and Multiple Level V Studies)

The use of calcium sulfate preparations mixed with local autograft, as a substitute for AICB, is an option for instrumented posterolateral fusions.

Grade I (Single Level V Study)

There is insufficient evidence to recommend for or against the use of a HA-glass/BMA composite as an autograft substitute for posterolateral fusion.

rhBMP-2: Interbody Fusion

Grade B (Multiple Level II Studies)

The use of rhBMP-2 as a substitute for AICB for ALIF with threaded interbody cages is an option due to similar fusion rates and clinical outcomes.

Grade C (Single Level II Study)

The use of rhBMP-2 as a substitute for AICB for single-level PLIF is an option due to similar fusion rates and clinical outcomes; however, formation of heterotopic bone has been observed.

Grade C (Single Level IV and Multiple Level V Studies)

The use of rhBMP-2 as a bone graft extender can be considered as an option when performing a TLIF procedure with a structural interbody graft.

Grade I (Single Level III Study)

There is insufficient evidence to make a recommendation regarding the use of rhBMP-2 as a supplement for stand-alone ALIF procedures using femoral ring allograft (single Level III study) or with a resorbable spacer when performing TLIF procedures (single Level V study).

rhBMP-2: Posterolateral Fusion

Grade B (Multiple Level II Studies)

The use of rhBMP-2 supplemented with 15% HA/85% β-TCP matrix as a substitute for AICB is an option in single-level posterolateral instrumented fusions given the consistent observation of comparable fusion rate and clinical outcomes.

Grade C (Single Level II and Single Level IV Studies)

The use of rhBMP-2 supplemented with graft extenders as an alternative to AICB is an option for single-level, instrumented posterolateral fusions in patients older than 60 years.

Grade C (Single Level III and Single Level V Studies)

The use of rhBMP-2 as a graft extender with either AICB or local bone is an option in patients undergoing either instrumented or noninstrumented posterolateral fusions.

Grade I

There is insufficient evidence to formulate a recommendation regarding the use of rhBMP-2/local bone as a substitute for AICB when performing revision posterolateral fusions (single Level III study) or the use of rhBMP-2/calcium-based extenders for single level posterolateral fusions in patients who smoke and elect to undergo surgery for lumbar spondylosis (single Level III study).

rhBMP-2: Complications

Grade C (Multiple Level IV and V Studies)

The use of rhBMP-2 as a graft option has been associated with a unique constellation of complications that the surgeon should be aware of when considering the use of this graft extender/substitute.

rhBMP-7

Grade C (Single Level II Study)

The use of rhBMP-7 when combined with local autograft as an alternative to AICB/local autograft is an option for single-level instrumented fusions based on equivalent clinical and radiographic outcomes. The use of rhBMP-7 has not been approved by the FDA for spinal fusions and currently requires a humanitarian device exemption.

Grade I (Conflicting Level II Studies)

No recommendation regarding the use of rhBMP-7/absorbable collagen sponge (ACS) as a substitute for AICB in posterolateral fusions can be made due to conflicting evidence from studies of equal strength.

Rationale

The objective of a lumbar fusion is to create an environment that will allow bone to form a solid osseous bridge across the involved spinal segments. Autologous iliac crest bone has been considered the gold standard because of its ideal graft characteristics, including osteoconduction, osteoinduction, and osteogenesis.[4,28,39] The harvesting of AICB, however, is commonly associated with increased postoperative pain, which may be underestimated by the treating surgeon.[31,55] Additional drawbacks of AICB include limited supply and increased operative time and blood loss.

Allograft bone, one of the original substitutes for AICB, may avoid some of these drawbacks; however, when used alone, it is commonly associated with an increased pseudarthrosis rate.[28] For this reason, and to avoid the morbidity of harvesting AICB, a great deal of time and expense has been dedicated to investigate and promote extenders and/or substitutes of AICB. Potential candidates include locally harvested autograft, calcium-phosphate salts, such as HA or β-TCP, and DBM. However, no material has received more attention and generated more controversy than the family of BMPs. There are numerous papers that demonstrate the fusion potential of BMPs;[2,7,18,19] however, complications associated with their use have been reported.[29,47,50] Whether the benefits of BMPs justify the costs remains to be determined. Possibly more alarming than the potential complications and costs have been questions related to bias and conflict of interest associated with the reporting of results from trials investigating the potential of BMPs.[9] This escalating controversy prompted the editors of *The Spine Journal* to dedicate the June 2011 issue to concerns regarding the use of BMPs in spinal fusion surgery.

The objective of this update is to build upon the previous recommendations formulated in the original guidelines publication.[48] A review of the recent medical literature was conducted to determine the utility of these materials with respect to their clinical efficacy, fusion potential, and complication risk. It is beyond the scope of the current update to comment on cost utility of these materials or the ethics of investigational reporting.

Search Criteria

A computerized search of the National Library of Medicine MEDLINE database, utilizing the online search engine PubMed, was conducted from 2003 through December 2011 utilizing the following search terms ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ((("Bone Substitutes"[MeSH] OR "Calcium Phosphates"[MeSH]) OR "Hydroxyapatites"[MeSH]) OR "Bone Morphogenetic Proteins"[MeSH]) AND (("2003"[PDAT]: "3000"[PDAT]) AND "humans"[MeSH Terms] AND English[lang]). The search was limited to the English language and human subjects and yielded a total of 151 papers. The titles and abstracts of these articles were reviewed and those specifically investigating the fusion potential, clinical efficacy, and potential complications of bone graft substitutes and extenders were selected. Of these papers a secondary review of the bibliographies was conducted to identify any additional relevant papers. A total of 79 articles were selected and reviewed in detail. Studies supporting similar conclusions of equivalent strength were grouped together. Those providing the best evidence from these compilations were included in the evidentiary tables. A detailed description of high-level studies or a representative of lower-level studies of similar conclusions serve as the scientific foundation for this update.

Scientific Foundation

Demineralized Bone Matrix

Since the publication by Urist, the osteoinductive properties of demineralized bone matrix (DBM) have been well recognized and extensively studied as both a substitute and extender of autograft bone.[57] Cammisa et al. conducted a multicenter, prospective, controlled trial to investigate the potential of DBM as a graft extender for AICB when performing a posterolateral instrumented lumbar fusion (Table 1).[8] One hundred twenty patients with a variety of degenerative disorders were enrolled and underwent up to a 3-level lumbar fusion. An independent, blinded radiologist, utilizing static and dynamic radiographs, performed fusion assessment at 3, 6, 12, 18, and 24 months. The clinical outcome of these patients was not recorded. All patients served as his/her own control receiving AICB within one intertransverse space and an equal volume of DBM/AICB to the contralateral intertransverse space. The follow-up rate at 24 months was 68%. A comparable fusion rate was observed on both sides (52% with DBM/AICB and 54% with AICB). Seventy-five percent of patients demonstrated fusion on both sides. Based on these observations the authors concluded that DBM could serve as an effective graft extender, decreasing the amount of autograft required and potentially reducing the risk and severity of donor site morbidity. Due to the utilization of internal controls, one cannot exclude a possible interaction between the investigational and control groups that could affect the outcome. This is therefore considered a case series when determining baseline level of evidence. Additional limitations include a heterogeneous patient population with respect to presenting diagnosis and the inclusion of a variety of fusions methods, including various interbody techniques. A large percentage of patients were lost to follow-up at 24 months. In the presence of pedicle screw stabilization, assessment of fusion with plain radiographs may be compromised. Secondary to these limitations, the case series was downgraded to Level V evidence in support of DBM as a graft extender for AICB in posterolateral lumbar fusion.

Schizas et al. conducted a pilot study comparing the clinical and radiographic outcome of patients undergoing 1- and 2-level posterolateral instrumented lumbar fusion using a novel DBM as a graft extender for autograft (Table 1).[51] Fifty-nine consecutive patients were divided into the 2 treatment groups; 33 received DBM mixed with autograft/BMA and 26 received only autograft. Fusion assessment

TABLE 1: Demineralized bone matrix: summary of evidence*

Authors & Year	Level of Evidence	Description of Study	Comments
Schizas et al., 2008	III	The objective of this pilot cohort study was to compare the clinical & radiographic performance of a novel DBM as a graft extender for autograft in 1- & 2-level posterolat instrumented fusions. 59 consecutive pts were divided into the 2 treatment groups: 33 received DBM mixed w/ autograft/BMA (treatment) & 26 received only autograft (control). Fusion assessment was blinded & determined w/ plain radiographs. Validated outcome measures were used to determine clinical status. At 12 mos after surgery, the fusion rate for the interventional group was 69.7% & that for the control group was 76.9% ($p = 0.57$). There was no difference in the clinical outcome btwn groups. The authors concluded that DBM is a safe & effective graft extender for 1- & 2-level posterolat fusions.	This study is limited by the relatively small study population w/ varying presenting diagnoses. Inadequate baseline demographic data are provided. The authors fail to standardize the DBM/autograft composite. Due to these limitations this study provides Level III evidence regarding the efficacy of DBM to act as a graft extender.
Cammisa et al., 2004	V	The authors conducted a multicenter prospective controlled trial to determine the efficacy of DBM as a graft extender for posterolat instrumented lumbar fusions. 120 pts were enrolled w/ varying degenerative diagnoses & underwent up to 3-level lumbar fusions. Radiographic assessment of fusion was performed at 3, 6, 12, 18, & 24 mos by independent blinded radiologists using static & dynamic radiographs. No clinical outcome measures were used. Each pt served as his/her own control & received equal vols of either AICB or a DBM/AICB composite to the intertransverse spaces. The follow-up rate at 24 mos was 68%. The fusion rate for the DBM/AICB composite side was 52% & 54% on the control side, w/ a 75% agreement btwn the 2 sides. Based on these observations the authors concluded that DBM can serve as an effective graft extender, decreasing the amount of autograft required & potentially reducing the risk & severity of donor site complications.	This study benefits from a relatively large no. of pts who were followed for an extended time; however, this study is considered a case series since control & treatment cohorts were conducted w/in the same pt & therefore possible interaction btwn the investigational & control sites cannot be excluded. A significant loss to follow-up rate (32%) was encountered. A variety of diagnoses were included & the surgical technique was not standardized. The use of plain radiographs may be an inadequate means of fusion assessment in the presence of pedicle screw instrumentation. The authors failed to include clinical outcome data. Due to these limitations the study was downgraded to Level V in support of DBM as a graft extender for AICB in pts undergoing instrumented posterolat fusions.

* AICB = autologous iliac bone graft; BMA = bone marrow aspirate; DBM = demineralized bone matrix; pt = patient.

was performed 12 months after surgery by a blinded independent observer utilizing plain radiographs. Validated outcome measures, including the Oswestry Disability Index (ODI) and visual analog scale (VAS) pain score, were used to determine clinical status. The fusion rate for the interventional group was 69.7% and the control group 76.9% ($p = 0.57$). There was no difference in the clinical outcome between the 2 groups. The authors concluded that DBM is a safe and effective graft extender for single- and 2-level posterolateral fusions. The relatively small number of patients, varying diagnoses, inadequate baseline demographic data, and short clinical follow-up all limit the conclusions of this study. The authors failed to standardized the volume and type of autograft used, whether AICB or locally harvested autograft. Due to these limitations the study is considered to provide only Level III evidence in support of DBM as a graft extender.

Calcium Phosphate Salts

This class of graft extenders and substitutes consists of calcium phosphate salts of varying composition that provide a lattice framework for in growth of new bone. These materials provide an osteoconductive matrix, having little if any osteoinductive or osteogenic properties. Examples include β-TCP, HA, and coral-based materials (Table 2).

β-Tricalcium Phosphate. Dai and Jiang performed a prospective, randomized, controlled trial to determine the efficacy of β-TCP as a bone graft substitute for AICB in single-level posterolateral instrumented fusions for degenerative spinal stenosis (Table 2).[16] Sixty-two patients were randomized to one of 2 cohorts, receiving either β-TCP ($n = 32$) or AICB ($n = 30$), both supplemented with local autograft. An independent observer assessed clinical outcome with validated outcomes measures, including mJOA, 36-Item Short Form Health Survey (SF-36), and VAS for donor site pain, at 6 weeks and 3, 6, 12, 24 and 36 months after surgery. Two independent observers evaluated plain radiographs to assess fusion status at 3, 6, 12, 24, and 36 months after surgery. No patients were lost to follow-up at 36 months. The reported fusion rate for all study participants at 36 months was reported to be 100%, and significant improvement in the clinical outcome was observed in all patients. No significant differences were observed between the treatment cohort and control group with respect to fusion rate or clinical outcome. A postoperative hematoma was reported in 3 patients undergoing harvest of AICB. All patients from the AICB group reported donor site pain after surgery, and only 20% reported no pain at 6 weeks after the operation. In some patients (data not provided) pain was still present 36 months after surgery. The authors concluded that β-TCP, supplemented with local autograft, could serve as an effective substitute for AICB when local bone is insufficient. This was a well-designed and executed trial of a homogeneous group of patients with excellent clinical follow-up over an extended time. There were several limitations, however, including a failure to describe the randomization process and failure to use a disease-specific outcome measure. The utilization of plain radiographs to determine fusion status may also be considered suboptimal, particularly in the presence of an instrumented fusion. Due to these limitations the study is downgraded to Level II evidence in support of utilizing β-TCP/local autograft as a substitute for AICB.

Hydroxyapatite. In 2005, Korovessis et al. conducted a prospective randomized controlled trial (RCT) to determine the fusion potential of coralline hydroxyapatite (CH) in multilevel, instrumented posterolateral lumbar fusions (Table 2).[38] Sixty patients were randomized to one of 3 cohorts: bilateral application of AICB, AICB on the left and CH/local bone/BMA on the right, and CH/local bone/BMA bilaterally. Validated outcome measures, including the ODI, VAS, Roland-Morris score, and SF-36, were obtained preoperatively and at 6, 12, 24, and 48 months after surgery. Two blinded, independent radiologists evaluated plain radiographs at 3, 6, 12, 24, and 48 months after surgery, supplemented with CT imaging at 12 and 24 months, to assess fusion status. Ninety-five percent of patients were available for follow-up at a minimum of 3 years. The fusion rate was 100% for all 3 groups at 1 year after surgery, based on CT and plain radiographs; however, in the CH/local bone/BMA cohort the fusion was limited to the facet joint and lamina. Reliability of radiographic assessment was adequate with an intraobserver and interobserver correlation coefficient (r) of 0.71 and 0.69, respectively. Improvement in all clinical outcome parameters was observed; however, no statistical analysis was performed to determine if any intergroup differences existed. The authors concluded that CH when combined with local bone/BMA is an appropriate AICB substitute when placed over the lamina and facet; however, this is inappropriate for intertransverse fusion. This was a comprehensive study of adequate design; however, the authors fail to provide adequate baseline demographic data to determine if pretreatment differences existed in the study groups. The authors avoided any limitations of utilizing an internal control by creating 2 additional study cohorts, those receiving only AICB and those receiving CH/local bone/BMA. Differences in clinical outcome are difficult to determine since no statistical analysis was performed. Due to these limitations the study is considered to provide Level II evidence in support of the authors' conclusions.

The objective of the single-center prospective cohort study conducted by Lee et al. was to determine the efficacy of a HA as a graft extender in instrumented posterolateral fusion (Table 2).[41] Thirty-three patients with varying diagnoses underwent either 1- or 2-level circumferential fusion with an HA/AICB mixture or AICB randomly applied to either the right or left intertransverse space. Equal volumes of graft material were used with the control arm receiving twice as much AICB as the investigational side, 6 versus 3 ml, respectively. Radiographs, obtained at 3, 6, and 12 months after surgery, and 3D thin-cut CT scans, obtained at 12 months, were independently reviewed to determine fusion status. Clinical outcome was not objectively recorded. The fusion status at 12 months was 86.7% in the investigational group and 88.9% in the control group. The volume of fusion was measured with CT and considered significantly greater on the investigational intertransverse space. The authors concluded that HA is a safe and effective graft extender for posterolateral lumbar fusion. A

TABLE 2: Calcium phosphate salts: summary of evidence*

Authors & Year	Level of Evidence	Description of Study	Comments
Dai & Jiang, 2008	II	The objective of this single-center prospective RCT was to determine the efficacy of β-TCP as a bone graft substitute for AICB in single-level posterolat instrumented fusions for degenerative spinal stenosis. 62 pts were randomized to one of 2 cohorts, receiving either β-TCP (n = 32) or AICB (n = 30), both supplemented w/ local autograft. Clinical data were collected by an independent observer at 6 wks & 3, 6, 12, 24, & 36 mos. Validated outcomes measures were utilized (mJOA, SF-36, & VAS for donor site pain). Fusion status was assessed by 2 independent observers evaluating plain radiographs at 3, 6, 12, 24, & 36 mos. The 36-mo follow-up rate was 100%. All study pts were considered to have achieved a solid arthrodesis & demonstrate significant improvement compared w/ their preoperative clinical status. There was no difference in clinical outcome btwn the 2 cohorts. 3 complications (postop hematomas) were associated w/ AICB harvesting, & donor site pain was still recorded in some pts at the 36-mo follow-up. The authors concluded that β-TCP can be used as an effective substitute for AICB when local bone is insufficient.	This is a relatively homogeneous, small group of pts w/ adequate information regarding baseline demographics. The follow-up interval is adequate w/ an excellent rate of follow-up. However, the randomization process was not described. Although validated outcome measures were used, the authors failed to utilize a disease-specific instrument to measure response to surgery. The method of fusion assessment, particularly in the presence of an instrumented spine, is suboptimal. It is not clear if the assessors of clinical & radiographic data were blinded to treatment. Due to these limitations, the study is downgraded to Level II evidence in support of β-TCP/local bone as a substitute for AICB/local bone in 1-level instrumented posterolat fusions.
Korovessis et al., 2005	II	This purpose of this prospective RCT was to determine the fusion potential of CHA in multilevel instrumented posterolat fusions. 60 pts were randomized into one of 3 groups: receiving AICB bilaterally, receiving AICB on the left & CHA/local bone/BMA on the right, & CHA/local bone/BMA bilaterally. Validated outcomes instruments (ODI, VAS, Roland-Morris, & SF-36) were obtained preoperatively & at 6, 12, 24, & 48 mos after surgery. Fusion assessment was performed w/ plain radiographs at 3, 6, 12, 24, & 48 mos postoperatively, supplemented w/ CT imaging at 12 & 24 mos. Images were evaluated by 2 independent radiologists blinded to the intervention. The follow-up rate was 95% at a minimum of 3 yrs. Reliability of radiographic assessment of plain radiographs was performed by repeating the evaluation after 3 wks. The fusion rate was 100% for all 3 groups at 1 yr after surgery, based on CT & plain radiographs; however, the fusion observed w/ application of CHA/local bone/BMA was limited to the facet joint & lamina w/o bridging bone in the intertransverse space. The intra- & interobserver coefficient values (r) for fusion assessment were 0.71 & 0.69, respectively. Improvement in all clinical outcome parameters was observed; however, no statistical analysis was performed to determine if any intergroup differences in clinical outcome occurred. The authors concluded that CHA when combined w/ local bone/BMA is an appropriate AICB substitute when placed over the lamina & facet but is inappropriate for intertransverse fusion.	This is a relatively small, heterogeneous population of pts w/ respect to preop diagnosis & no. of levels fused. The process of randomization was not adequately described. Although the authors state that the treatment groups were similar, adequate preop demographic data are not provided, & they failed to demonstrate equipoise btwn the treatment groups. The use of an internal control is suboptimal, since one cannot rule out interaction btwn the 2 intertransverse spaces; however, the authors controlled for this limitation by creating an isolated AICB & CHA/local bone/BMA groups. The authors failed to perform adequate statistical analysis of the clinical results. Due to these limitations the study was downgraded to Level II evidence supporting the efficacy of CHA combined w/ local bone/BMA as an AICB substitute.
Chang et al., 2008	IV	The authors performed a retrospective review of results from pts undergoing single-segment posterolat lumbar fusion using either local autograft expanded w/ calcium sulfate or iliac crest autograft. 115 pts were divided into treatment groups & evaluated over 1 yr. Similar fusion rates were observed, 92.3% for the treatment group (n = 66) & 92.9% in the control (n = 49), as well as clinical outcome, as determined by VAS score & ODI. The authors concluded that local autograft supplemented w/ calcium sulfate is a safe & effective alternative to autogenous iliac crest autograft.	The investigators failed to describe how pts were allocated into treatment groups, whether the evaluation of pts was blinded, performed an inadequate description of the results, inadequately described the statistical analysis of the data, & made statements w/in the conclusion that cannot be accounted for by their observations. This study provides Level IV evidence in support of local autograft w/ calcium sulfate as a substitute for AICB in single-segment posterolateral fusions.

(continued)

TABLE 2: Calcium phosphate salts: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Lee et al., 2009	V	The objective of this single-center prospective controlled cohort study was to determine the efficacy of a commercially available HA as a graft extender in instrumented posterolat fusion. 33 pts w/ varying diagnoses underwent either 1- or 2-level circumferential fusion w/ an HA/AICB mixture or AICB randomly applied to either the right or left posterolat space. Radiographs, performed at 3, 6, & 12 mos after surgery, & 3D thin-cut CT at 12 mos, were independently reviewed to determine fusion status. Clinical outcome was not objectively recorded. The fusion status of the posterolat space at 12 mos was 86.7% in the investigational group & 88.9% in the control group. The volume of fusion, measured w/ CT, was considered significantly greater in the investigational group. The authors concluded that HA granules are a safe & effective graft extender for posterolat fusion.	This study is considered a case series since control & treatment cohorts were conducted w/in the same pt, at the same level, & therefore cannot be considered independent variables. It is possible that an interaction occurred btwn the investigational & control sites that could have an impact on the outcome. The outcomes of the posterolat fusions may also be altered by the placement of the interbody graft; therefore, it is difficult to determine the true efficacy of HA as a graft extender w/in the posterolat space. The authors failed to include an objective measure of clinical outcome. The study population was small & heterogeneous w/ respect to presenting diagnosis. This case series is therefore downgraded to Level V evidence in support of HA as a graft extender.
Acharya et al., 2008	V	The aim of this study was to determine the efficacy of an HA–bioactive glass ceramic composite as a graft substitute for autologous bone in lumbar posterolat fusions. 24 consecutive pts undergoing posterolat instrumented lumbar fusion were entered into the study. The left intertransverse space received a standard mixture of BMA/HA-glass ceramic composite & the right side received an equal volume of locally harvested autograft. The primary outcome measure was graft consolidation as demonstrated on anteroposterior radiographs at 12 mos after surgery, evaluated by an independent orthopedic surgeon. At 1 yr after surgery, 22/24 pts were available for evaluation. The control side demonstrated a fusion rate of 73%; however, no definitive fusion was observed on the HA-glass/BMA composite side in any pt, w/ 77% of pts demonstrating complete graft resorption. The study was terminated at the request of the principal investigator. The authors concluded that there was no role for this HA-glass composite/BMA as a standalone graft substitute in posterolat fusion of the lumbar spine.	Although this study attempts to compare 2 treatment alternatives, a major limitation is the potential interaction btwn the control & interventional groups since both occurred in the same pt & therefore were not independent of each other. For this reason this study is considered a case series & not a comparative cohort study. This is a heterogeneous cohort, w/ respect to preop diagnosis & no. of levels fused. A suboptimal assessment of fusion was performed. The follow-up interval may be considered brief; however, 91% of enrolled pts were available. The dramatic difference in fusion outcome is noted. Due to these limitations the study is downgraded to Level V evidence against the use of HA-glass/BMA composite as a fusion substitute; however, it is included due to the dramatic difference in fusion rates.
Hsu et al., 2005	V	The purpose of this prospective controlled cohort study was to determine the effectiveness of local harvested autograft & CHA as graft extenders/substitutes for AICB w/ instrumented posterolat fusions. 58 pts were divided into 3 groups receiving different graft mixture in the right intertransverse space, AICB/local bone (n = 20), AICB/CHA (n = 19), & CHA/local bone (n = 19). All pts received AICB in the left intertransverse space. Fusion assessment using radiographs was independently performed by 2 reviewers at 4, 6, & 12 mos after surgery. If uncertainty existed a CT was performed. The fusion rates between the investigational & control groups at 12 mos were 90% for AICB/local bone vs 95% control, 78.9% for AICB/CHA vs 84.2% control, & 57.9% for CHA/local bone vs 89.5% control. The only statistically significant difference in reported fusion rate was btwn AICB/local & CHA/local. The authors concluded that either local bone or CHA can serve as a graft extender w/ AICB; however, CHA/local bone was not an acceptable substitute for AICB w/ instrumented posterolat fusions.	There were a relatively small no. of pts w/in each study subgroup. Heterogeneity existed btwn study groups w/ respect to diagnosis & surgery & incomplete demographic data are provided. Use of the contralat intertransverse space as a control is suboptimal since one cannot rule out an interaction btwn the control & investigational sides w/in the same pt. The authors failed to perform statistical analysis btwn the investigational & control arms w/ in each study subgroup. Due to these limitations the study was downgraded to Level V evidence in support of CHA as a graft extender.

(continued)

TABLE 2: Calcium phosphate salts: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Chen et al., 2005	V	The authors performed a prospective controlled cohort study to determine the effectiveness of artificial calcium sulfate mixed w/ local autograft vs AICB for fusion formation in 1- or 2-level instrumented posterolat fusions. Btwn September 2000 & September 2001, 74 consecutive pts were enrolled w/ a mean follow-up of 32.5 mos. 39 pts underwent a 1-level fusion & 35 a 2-level fusion. Each pt received AICB to one side & a mixture of calcium sulfate/local autograft to the other posterolat gutter. 2 independent blinded orthopedic surgeons evaluated radiographs for bridging intertransverse bone at 3-mo intervals for the 1st year & then on an annual basis. The protocol for clinical evaluation was not well defined, but it appears that ODIs were obtained in a retrospective fashion. The fusion rate for 1-level procedures was 87.2% on the investigational side & 89.7% on the control side ($p = 1.0$). For 2-level procedures the fusion rate was 82.9% for the investigational side & 85.7% for the control side ($p = 1.0$). Clinical outcome was judged as excellent or good in 78.3%. The authors concluded that calcium sulfate/local bone is as effective as AICB in achieving a solid fusion for instrumented posterolat fusions.	This is a heterogeneous study population w/ respect to presenting diagnosis & the authors fail to provide adequate baseline demographic data. Since each pt acted as his/her own control, it is impossible to rule out a possible interaction btwn the investigational & control sides. Application of graft was not randomized; therefore bias cannot be excluded w/ respect to fusion bed preparation. The method of fusion assessment was subjective. This study is downgraded to Level V evidence in support of calcium sulfate/local autograft as a graft substitute for AICB.

* AICB = autologous iliac bone graft; β-TCP = β-tricalcium phosphate; BMA = bone marrow aspirate; CHA = coralline hydroxyapatite; HA = hydroxyapatite; mJOA = modified Japanese Orthopaedic Association; ODI = Oswestry Disability Index; RCT = randomized controlled trial; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog score.

major limitation of the study is the use of the contralateral intertransverse space as the control arm in the study subjects. Interaction between treatment and control groups cannot be excluded; therefore, this is considered equivalent to a case series. Additional limitations include the small, heterogeneous population of patients and failure to include an objective assessment of clinical outcome. The study is considered to provide Level V evidence in support of HA acting as a graft extender for AICB.

Chang et al. performed a retrospective comparative study to determine if calcium sulfate mixed with local autograft could serve as an alternative to AICB for single-segment posterolateral fusions (Table 2).[12] One hundred fifteen patients were divided between treatment and control groups and were observed for longer than 1 year. Fusion was assessed through static and dynamic radiographs as well as reformatted CT images. Outcome was assessed utilizing VAS scores and ODIs. Similar fusion rates (92.3% for the treatment cohort [n = 66] and 92.9% for the control group [n = 49]) and clinical scores were observed between groups. This study benefitted from an accurate assessment of fusion and use of objective outcome measures, but it is limited by the investigators' failure to describe how patients were allocated to study cohorts, whether the evaluations were blinded, and an inadequate description of the statistical analysis. Because of these limitations, the study was downgraded to Level IV evidence in support of using a mixture of calcium sulfate and local autograft as a substitute for AICB.

Acharya et al. conducted a prospective cohort study to determine the efficacy of an HA-bioactive glass ceramic composite as a substitute for autologous bone (Table 2).[1] Twenty-four consecutive patients undergoing posterolateral instrumented lumbar fusion were entered into the study. Each patient served as his or her own control, with the left intertransverse space receiving a standard mixture of BMA and HA-glass ceramic composite and the right side receiving an equal volume of locally harvested autograft. Anteroposterior radiographs obtained 12 months after surgery were evaluated by an independent orthopedic surgeon and were used to assess fusion status. The follow-up rate at 1 year after surgery was 91%. A definitive fusion was demonstrated on the control side in 73% of patients; however, there was no clear evidence of a fusion on HA/BMA side for any patient, with 77% of patients demonstrating complete resorption of the HA-bioactive glass/BMA graft. Given the dramatic difference in outcome, the principal investigator terminated the study and the authors concluded that this HA-glass composite/BMA was ineffective as a graft substitute in posterolateral lumbar fusion. Although this study suffers from a relatively small, heterogeneous patient cohort, the outcome assessment was performed in an objective manner with an adequate follow-up rate at the end point of the study. However, the use of an internal control is considered inadequate, as one cannot exclude an interaction between the control and treatment sides. Such an interaction would likely bias in favor of the treatment arm. Despite this inference, this study was considered equivalent to a case series and not a comparative cohort study. Although the observed results demonstrate a significant difference in fusion potential, the study design and

limitations necessitate an assignment of Level V evidence that HA-bioactive glass when combined with BMA cannot serve as a substitute for autologous bone.

A number of lower-quality studies and case series, providing Level IV and V evidence, have also been published investigating the utility of various calcium-based graft extenders[21,22,32] or substitutes for AICB.[11,13,15,33,45] These materials were generally mixed with either local autograft or BMA (Table 2). Given the nature of the study design, control groups were poorly designed, historical, or absent; therefore, direct comparisons to AICB are difficult and lack appropriate validity. When compared with previous published results, investigators considered the fusion rates acceptable; however, many of these studies are unable to provide data with respect to the actual clinical benefit since baseline demographic data are not provided. Due to these limitations, these studies only demonstrate the feasibility of utilizing these calcium-based materials as graft extenders or substitutes.

Bone Morphogenetic Proteins

Since the introduction of BMPs by Marshall Urist in 1965,[57] the application of these fusion agents, intended to induce bone formation from surrounding tissue, has dramatically increased. In 2002, the US FDA granted premarket approval for the use of InFUSE (rhBMP-2, Medtronic Sofamor Danek) for single-level ALIF procedures from L-4 to S-1 when used in conjunction with the LT-CAGE Lumbar Tapered Fusion Device (Medtronic Sofamor Danek).[58] Under a humanitarian device exemption, the FDA subsequently approved osteogenic protein–1 (OP-1; rhBMP-7, Stryker) for revision posterolateral lumbar spine fusion, where harvesting of autograft was not possible or not expected to achieve solid arthrodesis.[60] Although the FDA had granted a similar approval for InFUSE/MASTERGRAFT (rhBMP-2, Medtronic Sofamor Danek) for revision of symptomatic, posterolateral lumbar pseudarthrosis, at the request of the sponsor, this device was withdrawn in 2010.[59] The use of these agents extends well beyond the FDA-approved applications, with approximately 85% of primary spine procedures utilizing BMP considered off label.[46] Although the off-label use of BMP for spine has met with radiographic and clinical success, concern has been raised due to reports of rare but significant neurological or structural complications following the use of BMPs, particularly with interbody fusions.[42,50] In addition, whether the routine use of BMPs is cost-effective has yet to be demonstrated. This uncertainty requires a careful evaluation of the literature investigating the various applications of the available BMPs.

rhBMP-2: Interbody Fusion. The utilization of rhBMP-2 as a substitute for AICB with threaded interbody cages for single-level ALIF procedures has been investigated in 2 randomized control trials.[5,6]

The larger of the 2 trials was previously evaluated in the original Lumbar Fusion Guidelines and was designated as a Level I study.[48] These higher-level studies were reevaluated for the purposes of this update since different criteria were used to determine levels of evidence and different recommendation grades formulated from the evidence. After reviewing the paper by Burkus et al.,[6] several limitations were identified including a failure to perform a power calculation to determine sample size, incomplete description of presenting demographic data (specifically no mention of comorbid medical conditions), and failure to perform an appropriate statistical analysis regarding outcomes between study cohorts (Table 3).

Burkus et al. also performed a prospective RCT in 42 patients to determine fusion progression of rhBMP-2 in a threaded titanium cage compared with AICB for single-level ALIF procedures.[5] The investigational cohort (n = 22) received rhBMP-2 and the control arm received AICB (n = 20). Fusion status was determined by 2 independent blinded radiologists evaluating both radiographs and CT images at 2 days and 6, 12, and 24 months after surgery. The fusion rate was 100% in the investigational cohort and 95% in the control group. The patients receiving rhBMP-2 demonstrated a greater average increase in bone density as demonstrated by Hounsfield units. The authors concluded that use of rhBMP-2 is associated with a high fusion rate and is a promising method to facilitate fusion in ALIF procedures. Given the dates of recruitment, these patients may have been included in a previous publication presented by the same authors and reviewed in the original Lumbar Fusion Guidelines.[6] This study focuses solely on the radiographic outcome of these patients without any inclusion of clinical data. The number of patients in each cohort is relatively small and varies with respect to presenting diagnosis. Although this is an RCT, inadequate baseline demographic data are included and the authors failed to determine appropriate sample size prior to initiating the study. This study therefore provides Level II evidence in support of rhBMP-2 as a substitute for AICB for single-level ALIF procedures with threaded interbody cages (Table 3).

Haid et al. conducted a multicenter prospective randomized controled study to investigate the clinical and radiographic outcomes of patients undergoing single-level PLIF utilizing either iliac crest bone graft (ICBG) or rhBMP-2/ACS.[29] Sixty-seven patients with single-level degenerative disc disease were randomized. Clinical outcome was assessed utilizing validated outcome measures at 6 weeks and at 3, 6, 12, and 24 months after surgery. Radiographs and CT scans were obtained at 6, 12, and 24 months after surgery. The follow-up rate at all time points was at least 89.6%. At 24 months after surgery the investigational group demonstrated a 92.3% fusion rate while only 77.8% were considered fused in the control group; this difference did not prove to be statistically significant. Significant clinical improvement was observed in both cohorts, with the investigational group demonstrated superior improvement in the back pain score at 24 months. Although considered to be clinically irrelevant, a significantly greater percentage of patients in the investigational group (71% vs 12%), had heterotopic bone formation posterior to the interbody cage. Sixty percent of controls continued to complain of donor site pain at 24 months. Due to concern regarding the significant increase in heterotopic bone formation, the authors terminated the study but concluded that these results were encouraging. Despite the lack of an observed consequence of this excessive bone formation, the authors elected not to con-

TABLE 3: rhBMP-2 in interbody fusion: summary of evidence*

Authors & Year	Level of Evidence	Description of Study	Comments
Haid et al., 2004	II	This multicenter prospective randomized nonblinded study was intended to evaluate the clinical & radiographic outcomes at 24 mos of pts undergoing 1-level PLIF using either ICBG or rhBMP-2/ACS. 67 pts w/ 1-level degenerative disc disease were randomized to study cohorts. Clinical outcome was assessed using validated outcome measures at 6 wks & 3, 6, 12, & 24 mos after surgery. Radiographs & CT scans were obtained at 6, 12, & 24 mos after surgery. The follow-up rate at all time points was at least 89.6%. The fusion rates at 24 mos were 92.3% & 77.8% in the investigational & control groups, respectively; the difference was not statistically significant. A significantly greater percentage of pts in the investigational group (71% vs 12%) had heterotopic bone formation posterior to the interbody cage. All clinical parameters improved significantly in both study cohorts compared w/ preop status. The improvement in back pain score was significantly greater in the treatment cohort vs controls at 24 mos. 60% of controls continued to complain of donor site pain at 24 mos. The authors concluded that the results were encouraging regarding the use of rhBMP-2 through the posterior interbody approach; however, further studies incorporating more refined surgical technique are required.	This study is limited by the nonblinded assessment of clinical outcome & relatively small pt cohorts. The process of randomization is not adequately described. The study was terminated due to concern regarding heterotopic bone formation posterior to the interbody cages. Although this proved to be clinically insignificant, the study was not restarted since the use of stand-alone PLIF cages had fallen out of favor. This study did provide Level II evidence in support of rhBMP-2 as a substitute for AICB in single-level PLIF w/ threaded cages.
Burkus et al., 2003	II	The authors conducted a prospective randomized controlled trial to determine fusion potential of rhBMP-2 in a threaded titanium cage compared w/ AICB for 1-level ALIF procedures. Btwn August 1998 & March 1999, 45 pts were randomized, w/ 42 (93%) available for follow-up at 24 mos after surgery. There were 22 pts in the investigational cohort & 20 in the control arm. 2 independent blinded radiologists evaluated both radiographs & CT images at 2 days & 6, 12, & 24 mos after surgery. Based on radiograph & CT criteria, the fusion rate was 100% in the investigational cohort & 95% in the control group. The pts receiving rhBMP-2 demonstrated a greater average increase in bone density as demonstrated by Hounsfield units; however, the rate of follow-up w/ respect to these data is not consistent w/ the overall follow-up rate. The authors concluded that use of rhBMP-2 is associated w/ a high fusion rate & is a promising method to facilitate fusion in ALIF procedures.	This study is limited by the relatively small pt population that varies w/ respect to presenting diagnosis. The authors fail to provide adequate baseline demographic data. The description of statistical analysis is limited & no power calculations were performed to determine sample size. Due to these limitations this study was downgraded to Level II evidence that rhBMP-2 can serve as a substitute for AICB for 1-level stand-alone ALIF procedure w/ threaded cages.
Burkus et al., 2002	II	The objective of this multicentered prospective randomized trial was to compare the efficacy of rhBMP-2 vs AICB for 1-level ALIF procedures using tapered cylindrical cages. 279 pts w/ symptomatic 1-level degenerative disease were enrolled, randomized in a 1:1 fashion to either receive rhBMP-2 (n = 143) or AICB (n = 136). Radiographic outcome was independently assessed using static & dynamic radiographs & CT scans at 6, 12, & 24 mos. Clinical assessment was conducted at 6 wks & 3, 6, 12, & 24 mos & included work status, neurological outcome, pain questionnaires, & ODIs. At 24 mos, the follow-up rate for both groups was >90%. Clinical outcome was comparable btwn groups. The rhBMP-2 group demonstrated a fusion rate of 94.5% at 24 mos while the fusion rate was for 88.7% of the control group. Donor site pain was reported in 32% of controls at 24 mos. The rate of secondary procedures was similar btwn groups. Based on these results the authors concluded that rhBMP-2 when used w/ a tapered cage is an effective technique for ALIF procedures & avoids the adverse effects of harvesting AICB.	This is a well-designed clinical trial & benefits from a large sample size, validated outcomes instruments, standardized surgical technique, comprehensive outcome analysis, & excellent follow-up rate. There are numerous limitations including a failure to perform power calculation to determine sample size, incomplete description of presenting demographics (no mention of comorbid medical factors), failure to describe inclusion/exclusion criteria, & failure to independently collect clinical data. The authors do not completely describe the distribution of complications btwn treatment & control groups. No statistics were performed to compare outcome btwn treatment & control groups. Due to the numerous minor limitations the study was downgraded to Level II evidence in support of rhBMP-2 as a substitute for AICB w/ 1-level ALIF procedures.

(continued)

TABLE 3: rhBMP-2 in interbody fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Slosar et al., 2007	III	The objective of this prospective cohort study was to determine if rhBMP-2 can safely accelerate allograft interbody fusions & reduce the no. of nonunions when compared w/ allograft alone. 75 pts w/ varying diagnoses & undergoing 1–3 level fusions were enrolled; 30 received an FRA w/ allograft croutons (control) & 45 received an FRA supplemented w/ rhBMP-2 (investigational); both groups received posterior pedicle stabilization w/o fusion. Fusion assessment, performed at 6, 12, & 24 mos, was blinded & performed using radiographs & CT images. Objective outcome instruments were used to determine clinical outcome; however, assessment was not blinded. At 24 mos after surgery, the control group follow-up rate was 97% & the investigational group rate was 96%. Statistically significant increase in fusion rate was observed at all time points for the investigational group (94%, 100%, & 100%) compared w/ controls (66%, 84%, & 89%). The investigational group demonstrated significantly better outcomes at 6 mos, but both groups demonstrated significant improvement at 12 & 24 mos compared w/ baseline. The authors concluded that allograft interbody fusion supplemented w/ rhBMP-2 significantly improved the fusion rate compared w/ allograft alone & the more rapid fusion observed w/ rhBMP-2 led to more rapid clinical improvement.	The study population includes pts w/ varying diagnoses & the authors fail to describe the distribution of these diagnoses btwn the 2 treatment groups. The potential impact of differences in the baseline demographic data & no. of levels operated btwn the cohorts cannot be determined since the authors failed to perform statistical analysis of these data. Although radiographic data were independently reviewed, it is not clear if the clinical assessment was blinded. Due to these limitations this study was downgraded to Level III w/ respect to rhBMP-2 as a supplement to anterior interbody fusion using femoral ring allograft & pedicle screw stabilization.
Mummaneni et al., 2004	IV	This retrospective cohort study was intended to compare the efficacy of rhBMP-2 w/ AICB when placed in an interbody spacer during a TLIF procedure. 44 pts underwent the TLIF procedure btwn September 2002 & December 2003. Follow-up data were available for 40 pts (90%). 19 pts received AICB in the interbody space, & 21 pts received rhBMP-2 supplemented w/ either AICB (n = 12) or local autograft (n = 9). Radiographic evaluation was performed w/ static & dynamic radiographs at 6-wk & 3-mo intervals. Clinical outcome was assessed at 3-mo intervals using the Prolo scale & VAS for donor site morbidity. The mean follow-up was 9 mos. The fusion rate in the AICB group was 95%. A 100% fusion rate was observed in pts receiving rhBMP-2 w/ at least 6 mos of follow-up; however, only 76% of pts from this cohort were available. At 6 mos after surgery, 58% of pts continued to complain of donor site pain. The authors concluded that the use of rhBMP-2 & local autograft is an excellent option when performing a TLIF procedure & avoids AICB donor site morbidity.	This is a heterogeneous, small cohort of pts that varied w/ respect to diagnosis & surgery. It is not clear how many pts underwent a supplemental posterolat fusion. The no. of pts is too small to perform any meaningful statistical analysis to determine differences w/in the treatment groups. The radiographic assessment of fusion was not clearly defined. Clinical outcome w/ respect to treatment efficacy was not conducted w/ validated outcomes measures. The follow-up interval was relatively short w/ a significant no. of rhBMP-2 pts not available. The evaluation of both clinical outcome & radiographs was not blinded. Due to these limitations the study is downgraded to Level IV evidence, supporting the role of rhBMP-2 as a supplement to interbody fusions when performing a TLIF.
Geibel et al., 2009	V	This case series describes the clinical & radiographic outcome of 48 pts undergoing 1- & 2-level PLIF & posterolat instrumented fusions w/ rhBMP-2. 37 (77.1%) underwent a 1-level procedure & 11 (22.9%) underwent a 2-level procedure. ODIs were obtained after surgery & CT images were obtained to assess fusion. The average follow-up was 17 mos. 2 independent radiologists evaluated the CT images & determined a fusion rate of 100%. Pt satisfaction, recorded as a willingness to repeat surgery, was observed in 89%. The authors concluded that rhBMP-2 can serve as a substitute for AICB & effectively & safely accomplish fusion through the PLIF approach.	This is a heterogeneous population of pts w/ respect to diagnosis & surgical procedure. No preop clinical data were recorded; therefore, the treatment effect of surgery cannot be determined. Due to these limitations the case series was downgraded to Level V evidence regarding the efficacy & safety of rhBMP-2 as a substitute for AICB w/ a PLIF.

(continued)

TABLE 3: rhBMP-2 in interbody fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Burkus et al., 2009	V	This report represents an update of the clinical & radiographic outcomes of pts from 2 separate studies undergoing stand-alone 1-level ALIF w/ tapered cages & rhBMP-2. No comparison w/ the control arm (pts receiving autologous iliac crest) was performed. At 6 yrs after the index procedure, 146 pts provided clinical follow-up, & radiographic evaluation was performed on 130 pts. A solid arthrodesis was identified in 98% of pts. There was no significant difference in any of the clinical outcome measures at 6 yrs compared w/ the values observed at 2 yrs after surgery. Significant improvement compared w/ preop scores was maintained at 6 yrs after surgery. 25 revision surgeries were performed over the 6-yr follow-up period & 7 btwn the 2- & 6-yr time points. The authors concluded that this technique was an effective method of obtaining an ALIF & maintaining long-term, significant clinical improvements.	The efficacy of rhBMP-2 as a substitute for autologous bone w/ tapered cages in 1-level ALIF was previously established w/ the prior studies. The results reported in this case series are limited by the significant no. of pts lost to follow-up, 48% dropout rate for clinical data, & 55% for radiographic evaluation & therefore was downgraded to Level V evidence. This case series does not retract or contribute significantly to the prior studies.
Rihn et al., 2009[49]	V	The objective of this retrospective review of a case series was to determine the clinical & radiographic outcomes of pts undergoing 1-level TLIF w/ rhBMP-2. 48 of 53 pts w/ varying diagnoses of lumbar degenerative disease were identified over a 2-yr interval. Static & dynamic radiographs were reviewed by an independent blinded spine surgeon at an average follow-up of 19.4 mos. Odom's criteria, pt satisfaction, & NRS of leg & back pain were obtained through telephone interview at an average follow-up of 27.4 mos. The fusion rate was 95.8%. 71% achieved excellent or good outcome, w/ 84% of pts satisfied w/ their outcome. Persistent back & leg pain was reported in 60.4% & 41.7%, respectively. An overall complication rate of 21.7% was reported, w/ approximately 25% attributed to the use of rhBMP-2. The authors concluded that adequate clinical & radiographic outcomes can be obtained.	This is a retrospective review of a case series of pts undergoing a 1-level TLIF. The study benefits from an independent assessment of fusion using appropriate radiographic methods. Validated outcome measures were incorporated to determine response to treatment. The heterogeneous pt population is considered a significant limitation, & therefore the study is downgraded to Level V evidence in support of rhBMP-2 used as an adjunct to interbody fusion.
Villavicencio et al., 2005	V	This retrospective case series describes the clinical & radiographic outcomes of 74 pts undergoing TLIF w/ allograft supplemented w/ rhBMP-2/ACS for a variety of degenerative disorders through a variety of surgical approaches. Pts were divided into subgroups based on the no. of levels fused & whether an open or minimally invasive approach was used. A single independent radiologist evaluated radiographs at 3, 6, 12, & 24 mos & CT images at 12 & 24 mos after surgery. Independent clinical assessment was performed at 12 mos after surgery utilizing MacNab criteria. 96% of pts were available at 12 mos after surgery. There were no adverse events directly related to the use of rhBMP-2. The fusion rate for the entire cohort was 100% by 10 mos after surgery. Improved clinical outcome was observed in all groups; however, those undergoing minimally invasive procedures tended to have a better outcome. The authors concluded that rhBMP-2 is a safe & effective bone graft extender when used in conjunction w/ allograft for the TLIF procedure.	This study suffers from significant heterogeneity w/ respect to diagnosis & surgical procedure. The no. of pts in each subgroup was relatively small, & incomplete description of demographic characteristics is provided. Nonvalidated outcome measures were used to assess clinical outcome. The study provides only Level V evidence in support of rhBMP-2 as a graft extender for allograft for TLIF procedures.
Lanman & Hopkins, 2004	V	This case series describes the clinical & radiographic results of 42 pts undergoing the TLIF procedure using rhBMP-2 & a bioresorbable interbody graft. Clinical & radiographic follow-up was conducted at 3, 6, & 12 mos after surgery using the ODI, plain radiographs, & CT images. The radiographic follow-up rate at 6 mos was 98% & at 12 mos was 26%. Preop clinical data were obtained in only 59% of pts w/ a follow-up rate of 92% at 6 mos & 36% at 12 mos. The fusion rate was 98% at 6 mos & 100% at 12 mos. Clinical improvement was observed at each study end point, although significance was not determined. There were no device-related complications. The authors conclude that these results indicate that the combination of rhBMP-2 w/ a resorbable spacer may be an appropriate alternative for interbody fusion & deserves further investigation.	This is a relatively poorly conducted case series w/ little information regarding the efficacy of rhBMP-2 used w/ a bioresorbable spacer. The study population is small & varies w/ respect to diagnosis, no. of levels fused, & op approach. An excessive no. of pts were lost during the follow-up period to make any judgments regarding clinical efficacy. It is not clear if the radiographic or clinical follow-up was blinded. The authors failed to perform any statistical analysis. This study provides only Level V evidence.

* ACS = absorbable collagen sponge; AICB = autologous iliac crest bone; ALIF = anterior lumbar interbody fusion; FRA = femoral ring allograft; ICBG = iliac crest bone graft; NRS = numeric rating scale; ODI = Oswestry Disability Index; PLIF = posterior lumbar interbody fusion; pt = patient; rhBMP-2 = recombinant human morphogenetic protein–2; TLIF = transforaminal lumbar interbody fusion.

tinue since the use of threaded titanium cages through the PLIF approach had fallen out of favor. The relatively small number of patients (< 50) in each cohort and the lack of a blinded clinical assessment limit conclusions formulated from this study. The study provides Level II evidence in support of using rhBMP-2 with threaded cages through the PLIF approach (Table 3).

Additional studies of lesser quality have explored the potential of rhBMP-2 as a graft extender for lumbar interbody fusion. Slosar et al. performed a prospective cohort study to determine the impact of rhBMP-2 on fusion rate and clinical outcome following ALIF with femoral ring allograft (FRA).[54] Seventy-five patients with varying diagnoses and undergoing up to a 3-level fusion were enrolled; 30 control patients (n = 30) received an FRA with allograft croutons and an investigational group (n = 45) received an FRA supplemented with rhBMP-2/ACS. A statistically significant increase in fusion rate was observed at all time points for the investigational group compared with controls. Both groups demonstrated significant clinical improvement at 12 and 24 months over baseline, but no significant difference was observed between treatment groups. A heterogeneous patient population with respect to presenting diagnosis and number of levels fused, an inadequate statistical analysis of potentially confounding baseline demographics, and failure to perform an independent, blinded clinical assessment requires that the study be downgraded to Level III in support of rhBMP-2 as an adjunct to FRA interbody fusion (Table 3).

Several retrospective cohort studies and case series have investigated the use of rhBMP-2 as a graft extender when performing TLIF with an interbody graft. Rihn et al. performed a retrospective review of 48 patients receiving rhBMP-2 during TLIF procedures and observed a fusion rate of 95.8% at the 2-year follow-up with 71% reporting excellent or good outcomes.[49] The complication rate was 21.7% with one-quarter of these complications attributed to the use of rhBMP-2. Villavicencio et al. reviewed the data from 74 patients undergoing either open or minimally invasive TLIF for varying diagnoses using rhBMP-2 and allograft.[67] The fusion rate for the entire cohort was 100%; however, a trend toward improved clinical outcome was observed for patients undergoing less invasive procedures. Mummaneni et al. conducted a retrospective cohort study intended to compare the efficacy of rhBMP-2 with AICB for TLIF.[44] Forty-four patients underwent a TLIF, with 40 patients (90%) available for a mean follow-up of 9 months. The control arm (n = 19) consisted of patients receiving AICB in an interbody spacer, while the investigational group (n = 21) received rhBMP-2 supplemented with either AICB (n = 12) or local autograft (n = 9). With at least 6 months of follow-up, the fusion rate in the AICB group was 95% and 100% in patients receiving rhBMP-2; however, only 76% of patients receiving rhBMP-2 were available for follow-up. At 6 months after surgery, 58% of patients continued to complain of donor site pain. The authors concluded that rhBMP-2 and local autograft is an excellent graft option and avoids donor site morbidity when performing a TLIF procedure. The relatively small, heterogeneous population of patients with respect to diagnosis and surgery performed limits these studies. Nonvalidated clinical outcome measures were used, and the method of fusion assessment is questionable given the presence of pedicle screw instrumentation. In the Mummaneni et al. study, neither the radiographic evaluation nor assessment of clinical outcome was performed in an independent, blinded fashion. Due to the baseline study designs and various limitations, these studies provide at best Level IV or V evidence in support of rhBMP-2 as a supplement for interbody fusion through the TLIF approach (Table 3). Additional case series have been published exploring the potential of rhBMP-2 as a graft extender or substitute.[7,24,40] Burkus et al. published a long-term clinical and radiographic companion study to their previous published report of patient undergoing single level ALIF procedures with stand-alone tapered cages and rhBMP-2/ACS.[7] No significant difference in outcome at 6 years was observed when compared with the previously published data obtained at 2 years after surgery. Geibel et al. reported a 100% fusion rate with an 89% patient satisfaction rate in 48 patients undergoing 1- and 2-level instrumented PLIF with rhBMP-2 and posterolateral fusion.[24] Lanman and Hopkins published the only case series investigating the use of rhBMP-2 in conjunction with a bioresorbable cage.[40] This study was limited by 64% of patients lost to follow-up at 12 months after surgery, compromising any attempt at a meaningful interpretation of the data. Since these studies are all case series with limitations, at best they provide only Level V evidence.

rhBMP-2: Posterolateral Fusion. In 2006, Dimar et al. reported the 24-month radiographic and clinical results of patients enrolled in an FDA investigational device exemption (IDE) study comparing rhBMP-2 combined with a compression-resistant matrix (CRM; bovine type I collagen carrier containing 15% HA and 85% β-TCP) with AICB in instrumented posterolateral fusions (Table 4).[19] Ninety-eight of 150 randomized patients were available for review. Clinical outcome, assessed using validated outcomes instruments (SF-36, ODI, and back/leg pain scores), was performed at 6 weeks and 3, 6, 12, and 24 months. Independent assessment of radiographs and CT images was performed at 6, 12, and 24 months. Operative parameters, including the surgical time and blood loss, were significantly less in the rhBMP-2/CRM cohort. Both groups demonstrated a significant clinical improvement compared with baseline, but not between treatment groups. The rhBMP-2/CRM cohort demonstrated a statistically higher fusion rate, 90.6% compared with 73.3%. At final follow-up, 16% of patients in the AICB cohort continued to complain of donor site pain. The authors concluded that rhBMP-2/CRM demonstrated similar clinical outcomes and improved fusion rates compared with AICB for single-level instrumented posterolateral fusions. Thirty-five percent of patients from the original cohort of randomized patients were lost to follow-up. This study included a heterogeneous patient population with respect to diagnosis. Due to these limitations the study was considered to provide level II evidence in support of rhBMP-2/CRM as a substitute for AICB.

Dimar et al. later reported the 2-year radiographic and clinical outcomes of a multicenter prospective randomized controlled IDE trial to investigate the use of rhBMP-2

TABLE 4: rhBMP-2 in posterolateral fusion: summary of evidence*

Authors & Year	Level of Evidence	Description of Study	Comments
Dimar et al., 2009	II	The purpose of this study was to report the 2-yr radiographic & clinical outcomes of a multicenter, prospective, randomized controlled IDE trial to investigate the use of rhBMP-2 matrix (bovine type I collagen carrier containing 15% HA & 85% β-TCP) as a substitute for autologous iliac crest for 1-level posterolat instrumented fusions. Well-described radiographic criteria were used to assess fusion in a blinded fashion, & validated outcome measures were completed to determine clinical outcome. Clinical follow-up was performed at 6 wks & at 3, 6, 12, & 24 mos & radiographic follow-up at 6, 12, & 24 mos. 463 pts were randomized btwn the treatment cohorts, w/ a 2-year follow-up rate of 89%. The control group demonstrated significantly longer op times & blood loss. The clinical outcome measures improved significantly compared w/ preop scores in both cohorts w/ no significant difference noted btwn treatment groups. 60% of the control group continued to complain of donor site pain at the 24-mo follow-up evaluation. The fusion rate, based on CT assessment, demonstrated a statistically significant difference in fusion rates at all times points; the final fusion rate was 96% in the rhBMP-2 group & 89% in the control group. There was no significant difference in adverse events except that the control group suffered 17 graft site–related events. The authors concluded that the use of rhBMP-2 matrix improved op parameters, achieved a higher fusion rate & comparable clinical outcomes, & therefore can eliminate the need for autologous iliac crest bone for 1-level posterolat instrumented fusions.	This is a well-designed trial. The authors failed to perform a power analysis to determine sample size & failed to account for pts lost to follow-up. It is not clear if the clinical assessment was blinded; however, given that the outcome measures were pt self-assessment instruments, this fact does not detract significantly from the observations & conclusions of the study. Due to these limitations the study was downgraded to Level II evidence to support the use of rhBMP-2 as a substitute for AICB in 1-level posterolat fusions.
Dawson et al., 2009	II	The purpose of this multicenter prospective randomized controlled trial was to investigate the use of rhBMP-2 on an ACS reinforced w/ 15% HA/85% TCP ceramic granules as a replacement for iliac crest autograft in posterolat instrumented fusions. Well-described radiographic criteria were used to assess fusion in a blinded fashion, & validated outcome measures were completed to determine clinical outcome. Overall success was determined by combining the results of these measures. Clinical follow-up was performed at 3, 6, 12, & 24 mos & radiographic follow-up at 6, 12, & 24 mos. 50 pts were randomized to the cohorts w/ a 24-mo follow-up rate of 88% in the treatment group & 86% in the control group. Improvements were observed in all clinical outcome measures in both groups w/ a trend toward superior improvement in the experimental group. At each time point, there was a trend for more successful fusion in the treatment group, w/ final fusion rates of 95% in the investigational group & 70% in the control group. No difference in the radiographic or clinical follow-up proved to be statistically significant. The authors concluded that the combination of rhBMP-2 & HA/TCP may be an effective alternative to autologous bone graft for 1-level posterolat instrumented fusions.	This is a well-designed study but was downgraded to Level II evidence due to the relatively small pt cohorts (<50 pts/cohort), incomplete description of baseline pt demographic data, failure to describe if the clinical assessment was blinded, & an overall outcome measure that has not been validated.

(continued)

TABLE 4: rhBMP-2 in posterolateral fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Glassman et al., 2008	II	The purpose of this prospective RCT was to compare the clinical & radiographic outcomes of patents >60 yrs undergoing instrumented posterolat fusions w/ either rhBMP-2/ACS or AICB. Various bone graft extenders were applied at the discretion of the surgeon in both cohorts. Validated outcome measures (SF-36, ODI, & NRS for back & leg pain) were administered. 2-yr follow-up data were collected in 94% of the pts (49 in the rhBMP-2 cohort & 51 in the AICB group). Baseline NRS leg pain was reported w/ greater frequency in the rhBMP-2/ACS cohort ($p = 0.031$); however, there were no other differences in baseline demographics. There was a statistically greater no. of complications in the AICB cohort (20 vs 8, $p = 0.014$), although none of the complications were directly attributed to either the harvest of AICB or the use of rhBMP-2/ACS. Statistically significant improvement was observed in all clinical outcome measures in both cohorts compared w/ baseline, although none of the differences btwn the cohorts was significant. CT images were obtained in 99 pts at 24 mos after surgery (93% follow-up rate). The fusion rate in the rhBMP-2/ACS cohort was 86.3% & 70.8% in the AICB group. The average CT grade was significantly higher in the rhBMP-2/ACS group (4.3 vs 3.8 [$p = 0.03$]). Revision surgery for nonunion was required in 1 pt in the rhBMP-2 cohort & 5 in the AICB group. An estimation of total cost over the 2 yrs was not significantly different btwn the 2 groups ($42,574 for the AICB cohort & $40,131 for the rhBMP-2/ACS cohort). The authors concluded that the study provided Level I evidence supporting the use of rhBMP-2/ACS as an AICB replacement for lumbar fusion in the older pt.	This is a heterogeneous population of pts w/ respect to the presenting diagnosis, level of involvement, & no. of levels included in the surgery. No power analysis was performed to determine sample size. The surgical procedure was not standardized, w/ bone graft extenders added to both groups at the discretion of the surgeon. It is not clear if the "grading" scheme of fusion assessment has been validated. No statistics were performed w/ respect to rate of fusion. Due to these limitations the study was downgraded to Level II evidence in support of rhBMP-2 as a substitute for AICB when supplemented w/ a graft extender for posterolat instrumented fusions in pts >60 yrs.
Dimar et al., 2006	II	The purpose of this prospective randomized nonblinded report was to present the 24-mo radiographic & clinical results of pts enrolled in an FDA IDE study undergoing 1-level instrumented posterolat fusion w/ either rhBMP-2/CRM or AICB. 98 of 150 pts, a follow-up rate of 65%, were available at 24 mos after surgery. Clinical assessment w/ validated outcomes instruments (SF-36, ODI, & back/leg pain scores) was performed at 6 wks & 3, 6, 12, & 24 mos. A blinded radiologist & 2 orthopedic surgeons independently evaluated radiographs & CT images at 6, 12, & 24 mos. The surgical time & blood loss were significantly less in the rhBMP-2/CRM cohort. Significant improvements in all clinical outcome measures compared w/ baseline values were observed in both groups, but no difference existed btwn groups. A significantly higher fusion rate was observed in the rhBMP-2/CRM cohort (90.6% vs 73.3%). 16% of pts in the AICB cohort continued to complain of donor site pain at 24 mos. The authors concluded that rhBMP-2/CRM demonstrated similar clinical results & improved fusion rates as AICB for 1-level instrumented posterolat fusions.	The heterogeneous pt population w/ respect to diagnosis limits this study & the significant loss to follow-up at the 2-yr study end point. Validated clinical outcomes were utilized & an effective method of fusion assessment was performed in a blinded fashion. Due to the significant loss to follow-up the study is downgraded to Level II evidence in support of rhBMP-2/CRM as a substitute for AICB in 1-level posterolat instrumented fusions.

(continued)

TABLE 4: rhBMP-2 in posterolateral fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Taghavi et al., 2010	III	The objective of this retrospective comparative study was to determine the efficacy of rhBMP-2/local bone to either allograft or autograft in revision instrumented posterolat fusions. 62 pts w/ varying initial diagnoses were included w/ minimum follow-up of 2 yrs. Pts were divided into 3 groups: Group 1 (n = 24) received rhBMP-2, Group 2 (n = 18) BMA/allograft, & Group 3 (n = 20) autograft. All received supplemental local bone. Fusion assessment, through static & dynamic radiographs, was performed by 3 blinded independent reviewers w/ a diagnosis of nonunion based on either surgical exploration if revision was performed or radiographic findings. Clinical outcome was determined through VAS scores. Group 1 demonstrated a fusion rate of 100%, Group 2 demonstrated a 77.8% fusion rate, & Group 3 had a 100% fusion rate. Pts undergoing multilevel procedures w/ BMA/allograft demonstrated a statistically lower fusion rate. No difference in VAS scores was observed. The authors concluded that rhBMP-2 could be an appropriate alternative to AICB in revision posterolat fusion.	Although an adequate description of baseline demographics is provided w/ appropriate statistical analysis. there remains the possibility of selection bias due to the potential differences in presenting diagnoses that the authors do not describe. The authors included an adequate assessment of fusion w/ acceptable radiographic criteria & incorporated validated outcome measures. Due to the retrospective nature of the study design, but lack of significant limitations, the study is considered Level III evidence in support of rhBMP-2/local bone as an alternative to AICB for revision posterolat fusions.
Singh et al., 2006	III	This is a prospective case-matched cohort study to determine if rhBMP-2 enhances fusion rate w/in a shorter time interval for pts undergoing instrumented posterolat fusion using AICB. 52 pts presenting w/ stenosis & spondylolisthesis were evaluated: 41 received rhBMP-2/AICB/local bone & 11 received AICB/local bone. Fusion assessment was performed w/ reformatted CT images & evaluated in a blinded, independent manner by 2 surgeons & 1 radiologist. 2 pts were lost to follow-up at the 2-yr time point. The fusion rates in the rhBMP-2 & AICB cohorts were 97% & 77%, respectively. Pts receiving rhBMP-2 were judged to fuse faster & demonstrate more robust fusions. Fusion rate & quality of fusion proved to be statistically superior in the rhBMP-2 cohort. No complication was attributed to the use of rhBMP-2. The authors concluded that rhBMP-2 may serve as a safe & effective supplement to AICB for posterolat instrumented fusion.	Limitations of this study include the potential for selection bias since an inadequate description of baseline demographics is provided & the authors failed to evaluate these data for significant differences, e.g., it is not known if the no. of levels fused was comparable btwn treatment groups. The no. of pts w/in the control group is relatively small. It is not clear whether the surgical technique was standardized between cohorts, w/ no mention regarding the amount of AICB used. The value of the outcome parameters, in particular the subjective assessment of fusion quality, is of limited value & has not been shown to impact clinical outcome. Due to these limitations the study was downgraded to Level III evidence in support of rhBMP-2 as a graft extender for AICB in instrumented posterolat fusions.

(continued)

TABLE 4: rhBMP-2 in posterolateral fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description of Study	Comments
Glassman et al., 2007[27]	III	The objective of this retrospective review of data collected during a prospective, randomized, unblinded trial was to determine the influence of smoking on fusion rate & outcome of pts receiving either AICB or rhBMP-2 for single-level posterolateral fusions. All pts were evaluated at a minimum of 2 yrs after surgery. Clinical outcome was measured utilizing validated outcomes instruments (ODI, SF-36, back & leg pain scores) at 6 wks & 3, 6, 12, & 24 mos after surgery. An independent, blinded assessment of fusion status was performed w/ static & dynamic radiographs & CT imaging at 6, 12, & 24 mos after surgery. The records of 148 pts were reviewed, 42 smokers & 106 nonsmokers. The smokers were equally distributed btwn the 2 cohorts, w/ 55 nonsmokers in the rhBMP-2 cohort & 51 in the AICB group. Fusion rate at 24 mos based on radiographs was 100% in the rhBMP-2 nonsmokers, 95.2% in the rhBMP-2 smokers, 94.1% in the AICB nonsmokers, & 76.2% in the AICB smokers. A significant difference was observed btwn all smokers (85.7%) & all nonsmokers (97.2%) as well as btwn smokers & nonsmokers receiving AICB. Similar rates of fusion were determined w/ CT imaging. Clinical outcome was improved in all parameters in all 4 cohorts, w/o a significant difference in the degree of improvement btwn groups. When considered collectively, the nonsmokers consistently demonstrated better ODI & SF-36 scores. The authors concluded that rhBMP-2/CRM may enhance fusion rate in cigarette smokers undergoing single-level fusion & that smoking is detrimental to clinical outcome regardless of fusion status.	The retrospective nature of this study does limit inferences formulated from the results despite the fact that the data were collected during a prospective RCT. Since pts were not randomized w/ respect to smoking status, the benefit of the randomization process is nullified. Given the study design, this report provides Level III evidence regarding the impact of rhBMP-2/CRM on fusion rate in smokers & the impact of smoking on clinical outcome.
Lee et al., 2009	IV	The objective of this retrospective comparative study was to determine the efficacy of rhBMP-2 in pts >65 yrs undergoing an instrumented posterolat fusion. 127 pts w/ lumbar degenerative disease were divided into 3 groups: pts >65 yrs received both rhBMP-2 & allograft (Group A, n = 34), pts <65 yrs received rhBMP-2 & allograft (Group B, n = 52), & pts >65 yrs received autograft (Group C, n = 41). Fusion assessment was performed by an independent blinded radiologist using static & dynamic radiographs, supplemented w/ CT when fusion was questionable. Kirkaldy-Willis & VAS scores were used to determine clinical outcome. The average follow-up was ~36 mos, although VAS scores were obtained up to 24 mos after surgery. The fusion rates were 82.4% in Group A, 94.2% in Group B, & 78.1% in Group C. At 2 yrs, Group B demonstrated a statistically superior VAS score compared w/ Group A. Clinical outcome was not statistically different btwn the 3 groups. The authors concluded that rhBMP-2/allograft yields equivalent outcomes to AICB in pts >65 yrs, but is not able to overcome all negative comorbidities associated w/ age.	There are limited baseline demographic data provided, w/ an inadequate description of comorbidities that would affect fusion potential. Apparent differences exist w/ respect to several of these comorbidities, e.g., no. of levels fused, & the authors fail to determine if these differences were significant. The existence of selection bias therefore cannot be excluded. The surgical technique is not described. An adequate assessment of fusion was performed, although the authors include time to fusion, which is of questionable significance. Due to these limitations the study was downgraded to Level IV evidence in support of rhBMP-2/allograft as a substitute for AICB in pts >65 yrs undergoing 1-level, instrumented, posterolat fusions.
Hamilton et al., 2008	V	The authors performed a retrospective review of pts undergoing noninstrumented posterolat fusion using local harvested autograft supplemented w/ rhBMP-2/ACS. 47 of 55 pts, median age 68.2 years, were evaluated. Fusion assessment was independently performed using radiographs & CT scans at various time points up to 36 mos after surgery. An independent clinical evaluation using validated instruments was performed at least 29 mos & up to 36 mos after surgery. The fusion rate was 80%. Greater than 85% of pts demonstrated an improved clinical outcome & pain relief. The authors concluded that the use or rhBMP-2 as a supplement to noninstrumented posterolat fusions leads to improved pain & a comparable fusion rate in elderly pts.	This case series contains a heterogeneous study population w/ respect to fusion justification & no. of levels involved. Although the clinical outcome was performed by an independent assessor, pts were required to "recall" their pre- & postop pain & health status. The description of clinical results is difficult to interpret. CT evaluation was only performed in 85% of study pts. Due to these limitations, the case series is downgraded to Level V evidence that rhBMP-2 is an effective bone graft extender.

* ACS = absorbable collagen sponge; AICB = autologous iliac crest bone; β-TCP = β-tricalcium phosphate; BMA = bone marrow aspirate; CRM = compression-resistant matrix; HA = hydroxyapatite; IDE = investigational device exemption; NRS = numeric rating scale; ODI = Oswestry Disability Index; pt = patient; RCT = randomized controlled trial; rhBMP-2 = recombinant human bone morphogenetic protein–2; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale.

matrix as a substitute for AICB for single-level posterolateral instrumented fusion (Table 4).[20] Fusion assessment was performed in a blinded fashion, and clinical status was evaluated through validated outcome measures. Clinical follow-up was performed at 6 weeks and at 3, 6, 12, and 24 months and radiographic follow-up at 6, 12, and 24 months. Four hundred sixty-three patients were randomized, with a 2-year follow-up rate of 89%. Significantly longer operative times and greater blood loss was observed in the control group. There was no statistical difference in clinical outcome between groups, although both demonstrated significant improvement compared with baseline scores. Donor site pain was reported in 60% of the control group at final follow-up. Based on CT imaging, the interventional group demonstrated a statistically superior fusion rate at 6 months (79% with rhBMP-2 and 65% with AICB [p = 0.002]) and at 24 months [96% with rhBMP-2 and 89% with AICB [p = 0.014]). The overall rate of adverse events was statistically similar; however, 17 graft-related complications were recorded in the control group. The authors concluded that the use of rhBMP-2 matrix improved operative parameters, led to a higher fusion rate, and achieved comparable clinical outcomes to AICB and therefore can be considered an acceptable substitute for single-level posterolateral instrumented fusion. This was a well-designed and well-executed study. It is not clear if the clinical assessment was blinded; however, utilization of patient self-assessment questionnaires decreases the likelihood of bias with reporting. The authors did not perform an appropriate power analysis to determine sample size and failed to provide information regarding patients lost to follow-up. Due to these limitations the study was downgraded to Level II evidence in support of rhBMP-2 matrix as a substitute for AICB.

Dawson et al. conducted a multicenter prospective RCT to investigate if rhBMP-2/ACS supplemented with an HA/TCP extender could serve as an appropriate substitute for AICB in instrumented posterolateral fusions (Table 4).[17] Fifty patients were randomized; clinical follow-up was performed at 3, 6, 12, and 24 months, and radiographic follow-up at 6, 12, and 24 months. At 24 months, the follow-up rate was 88% in the treatment group and 86% in the control group. Both groups demonstrated improvements in all clinical outcome measures with the rhBMP-2 cohort demonstrating a trend toward better outcomes. The fusion rate was higher at all time points in the rhBMP-2 cohort, with final fusion rates of 95% in the investigational group and 70% in the control group. No difference in the radiographic or clinical outcome proved to be statistically significant. The authors concluded that the combination of rhBMP-2 and HA/TCP could be an effective alternative to AICB for single-level posterolateral instrumented fusions. The relatively small numbers of patients (< 50 patients per treatment arm), failure to provide adequate baseline demographic data, and utilization of a nonvalidated composite score to assess overall success are limitations of the study. The study was therefore considered to provide Level II evidence in support of rhBMP-2/ACS and HA/TCP as a graft substitute for instrumented posterolateral fusions.

Glassman et al. conducted a prospective RCT to compare the clinical and radiographic outcomes of 106 patients older than 60 years of age undergoing instrumented posterolateral fusions with either rhBMP-2/ACS or AICB (Table 4).[26] The method of grafting was not standardized, with various graft extenders added at the discretion of the surgeon. Clinical outcome was determined with validated outcome measures, including SF-36, ODI, and numeric rating scale (NRS) for back and leg pain. Computed tomography scans were used to assess fusion. At 24 months after surgery the clinical and radiographic follow-up was 94% and 93%, respectively. At baseline, the patients in the rhBMP-2 cohort reported leg pain with greater frequency (p = 0.031); there were no other differences in baseline demographics. The complication rate was significantly greater in the AICB cohort (20 vs 8, p = 0.014), although none of the complications were directly attributed to either the harvest of AICB or the use of rhBMP-2/ACS. Both cohorts demonstrated a statistically significant clinical improvement over baseline; however, there was no difference in clinical outcome between treatment groups. An 86.3% fusion rate was observed in the rhBMP-2/ACS cohort, compared with 70.8% in the AICB group. The authors provided a CT "grade" for the observed fusion with the rhBMP-2 cohort demonstrating a significantly higher score (4.3 vs 3.8 [p = 0.03]). Nonunion requiring revision was reported in 1 patient in the rhBMP-2 cohort and 5 in the AICB group. An estimation of total cost over 2 years was calculated, and the difference between the 2 groups was not significantly different ($42,574 for the AICB cohort and $40,131 for the rhBMP-2/ACS cohort). The authors concluded that the study provided Level I evidence supporting the use of rhBMP-2/ACS as an AICB replacement for lumbar fusion in the older patient. This study suffers from several limitations, including a heterogeneous patient cohort with respect to presenting diagnosis, failure to account for patients lost to follow-up, lack of a standard surgical protocol, questionable "grading" scheme to assess fusion, failure to determine sample size through a power analysis, and failure to perform an adequate statistical analysis. Due to these limitations the study was downgraded to Level II evidence in support of rhBMP-2 for patients older than 60 years of age undergoing posterolateral lumbar fusions.

Singh et al. published a prospective cohort study to compare outcome of patients receiving a mixture of rhBMP-2/local bone/AICB (n = 41) to those receiving only local bone/AICB (n = 11).[53] Fusion assessment was performed in an independent, blinded manner with reformatted CT images. The fusion rate with rhBMP-2 was 97% while the control cohort demonstrated a fusion rate of 77%. Those receiving rhBMP-2 were thought to achieve fusion faster and demonstrate a more robust fusion. However, this study failed to provided an adequate description of baseline demographics (for example, the number of levels fused in each group). This is a small and heterogeneous population of patients; it is not clear if the surgical procedure was standardized between cohorts, and no objective clinical outcomes were reported. The study was therefore downgraded to Level III evidence in support of utilizing rhBMP-2 as an extender.

Hamilton et al. published a retrospective case series of patients undergoing noninstrumented posterolat-

eral fusions utilizing local autograft supplemented with rhBMP-2.[30] An 80% fusion rate was observed; 85% of patients were felt to demonstrate clinical improvement. This study provides only Level V evidence in support of rhBMP-2 as an extender with local bone for noninstrumented fusions due to the heterogeneous patient population, failure to collect clinical outcomes in a prospective manner, and a radiographic follow-up rate of only 85%.

Taghavi et al. performed a retrospective cohort study to determine the efficacy of rhBMP-2/local bone to either allograft combined with bone marrow aspirate or autograft, in revision instrumented, posterolateral fusions. Sixty-two patients with varying diagnoses were included with a minimum follow-up of 2 years.[56] Patients were divided into 3 groups: Group 1 (n = 24) received rhBMP-2, Group 2 (n = 18) received BMA/allograft, and Group 3 (n = 20) received autograft. The exact source of autograft bone for Group 3 was not clearly defined. All 3 cohorts received supplemental local bone. Static and dynamic radiographs were used to assess fusion and were reviewed by 3 blinded independent reviewers with a diagnosis of nonunion based on either surgical exploration if revision performed or radiographic findings. Clinical outcome was determined through VAS scores. A fusion rate of 100% was observed for Groups 1 and 3; however, Group 2 demonstrated a 77.8% fusion rate. Patients undergoing multilevel procedures with BMA/allograft demonstrated a statistically lower fusion rate. No difference in VAS scores was observed. The authors concluded that rhBMP-2 could be an appropriate alternative to AICB in revision posterolateral fusion. Although this study was well executed with appropriate follow-up, validated outcome measures, and appropriate assessment of fusion, due to the retrospective nature of the study design it provides Level III evidence. As this was the only study identified to investigate the use of rhBMP-2 in revision surgery, evidence is insufficient to formulate a recommendation. Additional studies of similar or lesser quality, such as retrospective reviews or case series, promoting the use of rhBMP-2 for various clinical scenarios, such as in patients who use tobacco, have been published.[3,25,27,36] Due to an insufficient number of such studies no formal recommendations could be constructed regarding the use of rhBMP-2 under the specific clinical circumstances.

rhBMP-2: Complications. Between 2003 and 2007, the annual number of procedures utilizing BMPs increased by 4.3-fold (from 23,900 to 103,194 cases), with spinal fusions accounting for almost 93% of these cases.[46] Although it is difficult to deny a positive impact on fusion rate, surgeons must be aware of the potential risks and complications related to the use of BMPs, particularly since the majority of procedures would be considered off label.

Rihn et al. performed a single-center retrospective cohort study to specifically identify complications associated with the use of rhBMP-2 for single-level TLIF procedures and to determine if these complications differed compared with the use of AICB.[50] Between January of 2004 and May of 2007, 130 patients underwent a single-level TLIF using either AICB or rhBMP-2 (Table 5). One hundred nineteen patients were available for review, 33 receiving AICB and 86 receiving rhBMP-2, with an average radiographic follow-up of 19.1 months and an average clinical follow-up of 27.6 months. A combination of plain radiographs and CT images were used to assess fusion status. Those patients receiving AICB demonstrated a 96.5% fusion rate, and the fusion rate in the rhBMP-2 cohort was 97% (p = 0.09). The overall complication rate was higher in the autograft cohort (45.5% vs 29.1%), but the difference was not significant. Donor site morbidity was the most common complication associated with AICB, and postoperative radiculitis was more often observed in the rhBMP-2 cohort (14% vs 3% [p = 0.08]). A significant decrease in radiculitis was observed after 2006 (20.4% to 5.4% [p = 0.047]), following the utilization of a hydrogel sealant intended to shield the exiting root. Additional complications thought to be related to the use of rhBMP-2 included osteolysis and heterotopic bone formation. The authors concluded that the TLIF procedure, regardless of graft, is associated with a relatively high complication rate (33.6% for the entire cohort). Although rhBMP-2 eliminates donor site morbidity, the surgeon should be aware of additional complications, such as radiculitis, osteolysis, and heterotopic bone formation, that may be associated with its use. The authors failed to disclose if the assessment, clinical or radiographic, was performed in a blinded fashion. No validated outcome measures were used, and the fusion criteria were not adequately described. This study provides Level IV evidence supporting the use of rhBMP-2; however, more importantly, it highlights several of the more common complications thought to be associated with the interbody application of rhBMP-2.

Pradhan et al. observed an increased rate of graft resorption, fracture, or collapse in patients undergoing standalone ALIF with femoral ring allograft.[47] Lewandrowski et al. observed osteolysis of the vertebral endplate following minimally invasive TLIF and speculated that endplate violation during interbody decortication may have been a contributing factor.[42] Although not specific for lumbar procedures, Vaidya et al. observed a higher incidence of graft subsidence when rhBMP-2 was used with an allograft interbody spacer.[66] Joseph and Rampersaud observed a greater incidence of heterotopic bone formation following the use of rhBMP-2 for minimal access interbody lumbar fusion, but no clinical sequelae associated with this excessive bone formation were identified.[34] Mindea et al. observed a higher incidence of postoperative radiculitis of a nonstructural cause associated with the use of rhBMP-2 during minimal access TLIF procedures.[43] Garrett et al. reported on the formation of painful postoperative seromas following the use of rhBMP-2 during posterolateral fusions.[23] Finally, Carragee et al. identified a higher incidence of retrograde ejaculation in patients receiving rhBMP-2 during ALIF procedures.[10] There have been a number of additional retrospective reviews and case series that have corroborated the findings from these reports.[3,14,37,52]

Although a direct cause and effect relationship between the use of rhBMP-2 and these complications cannot be formulated based on these studies, the potential association should not be ignored. The surgeon should carefully consider the off-label utilization of rhBMP-2, or any osteobiologic, and make sure that the patient has been adequately informed regarding these risks.

TABLE 5: Complications and the use of rhBMP-2: summary of evidence*

Authors & Year	Level of Evidence	Description of Study	Comments
Rihn et al., 2009[50]	IV	This retrospective cohort study was intended to evaluate the complications associated w/ 1-level TLIF procedures & determine if complications differed btwn pts receiving AICB or rhBMP-2. 119 of 130 pts were available for follow-up, 33 receiving autograft & 86 rhBMP-2. After 2006, a hydrogel sealant was used in 37 pts to shield the nerve root from exposure to the rhBMP-2. Fusion assessment, through a combination of radiographs & CT scans, was performed on average at 19.1 mos & clinical follow-up at 27.6 mos after surgery. The fusion rate for the autologous cohort was 96.5% & it was 97% for the rhBMP-2 cohort ($p = 0.09$). A higher complication rate was observed in the autograft cohort (45.5% vs 29.1%), but the difference was not significant. Donor site morbidity was the most common complication in the autologous cohort. Postop radiculitis was more commonly observed in the rhBMP-2 cohort (14% vs 3% [$p = 0.08$]); however, this incidence significantly decreased after the application of the hydrogel sealant (20.4% to 5.4% [$p = 0.047$]). Additional complications thought to be related to the use of rhBMP-2 included osteolysis & heterotopic bone formation. The authors concluded that the TLIF procedure, regardless of graft, is associated w/ a relatively high complication rate (33.6% for the entire cohort), & although rhBMP-2 eliminates donor site morbidity, additional complications such as radiculitis, osteolysis, & heterotopic bone formation are associated w/ its use.	The study is limited due to its retrospective design & failure to describe if the assessment of complications & fusion status was performed in a blinded fashion. No validated outcome instruments were used. The criteria for fusion assessment were not adequately described, & it is not clear if similar methods were performed equally for both study groups. This retrospective cohort study is downgraded from Level III to Level IV w/ respect to the complications associated w/ the use of rhBMP-2 in TLIF procedures.

* AICB = autologous iliac crest bone; pt = patient; rhBMP-2 = recombinant human bone morphogenetic protein–2; TLIF = transforaminal lumbar interbody fusion.

rhBMP-7: Posterolateral Fusion. Osteogenic protein–1, also known as rhBMP-7, is another member of the transforming growth factor–β superfamily and plays a key role in osteoblast differentiation. Compared with rhBMP-2, there have been relatively few clinical studies investigating rhBMP-7 as an agent of spinal fusion, with the majority, if not all, focusing on posterolateral fusion techniques. The FDA has not approved the use of rhBMP-7 for spinal fusions. Currently, its use requires a humanitarian device exemption. There were no recommendations pertaining to the use of rhBMP-7 in the original guidelines publication (Table 6).[48]

Delawi et al. published a prospective multicenter RCT comparing the efficacy of local autograft supplemented with either rhBMP-7 or AICB in single-level instrumented posterolateral fusions for isthmic or degenerative spondylolisthesis.[18] Clinical and fusion assessment was performed at 6 weeks, and 3, 6, and 12 months after surgery. Fusion assessment through CT imaging was performed in a blinded fashion. Validated clinical measures (ODI and VAS) were used to determine response to surgery. The follow-up rate at 12 months was 89%. There was no statistical difference in the fusion rate between the groups (63% in the rhBMP-7 group and 67% in the AICB cohort). Both groups demonstrated clinical improvement compared with baseline scores, but there was no statistical difference regarding clinical outcome between groups. At 12 months after surgery, 64% of AICB patients complained of at least "mild" donor site pain (2.7 ± 2.8 VAS). No specific adverse event was related to AICB harvesting or use of rhBMP-7. The authors concluded that rhBMP-7 is an effective alternative to AICB for supplementing local autograft for single-level posterolateral fusions. The small sample size (< 50 patients), incomplete statistical analysis, and potential impact of differences in baseline characteristics requires that the study be downgraded to Level II evidence in support of utilizing rhBMP-7/local autograft as a substitute for AICB/local autograft in single-level instrumented posterolateral fusions.

In 2008, Vaccaro et al. conducted a multicenter prospective RCT to further investigate the safety and efficacy of rhBMP-7/ACS and to demonstrate noninferiority as a replacement for AICB for noninstrumented, single-level posterolateral fusion.[62] Three hundred thirty-five patients were randomized in a 2:1 fashion, but only 293 were treated (208 patients received rhBMP-7/ACS and 87 received AICB). Independent blinded clinical and radiographic evaluations were performed at 6 weeks and at 3, 6, 12, 24, and longer than 36 months utilizing validated outcome measures, including ODI, SF-36, and VAS, and radiographs. Fusion assessment after 36 months was supplemented with CT scans. The primary overall success was reported as a composite measure, intended for FDA submission, and required a 20% improvement in ODI, absence of treatment-emergent adverse events related to the device, absence of a decline in neurological status, and radiographic successful fusion. At 24 months with a follow-up rate of 87%, the investigational group did not achieve statistical equivalence with respect to the overall success rate compared with controls (38.7% compared with 49.4% [$p = 0.33$]). The investigational group demonstrated a

TABLE 6: rhBMP-7 in posterolateral fusion: summary of evidence*

Authors & Year	Level of Evidence	Description	Comments
Delawi et al., 2010	II	The objective of this pilot, prospective, randomized, controlled multicenter trial was to compare the efficacy of rhBMP-7 w/ AICB, both supplemented w/ local bone, in 1-level instrumented posterolat fusions for isthmic or degenerative spondylolisthesis. Clinical & fusion assessments were performed at 6 wks & 3, 6, & 12 mos after surgery. Fusion assessment through CT imaging was performed in a blinded fashion. Validated clinical measures (ODI & VAS) were used to determine response to surgery. The follow-up rate at 12 mos was 89%. There was no statistical difference in the fusion rate btwn the groups (63% in the rhBMP-7 group & 67% in the AICB cohort). Both groups demonstrated clinical improvement compared w/ baseline scores, but there was no statistical difference regarding clinical outcome btwn groups. At 12 mos after surgery, 64% of AICB pts complained of at least "mild" donor site pain (2.7 ± 2.8 VAS). No specific adverse event was related to AICB harvesting or use of rhBMP-7. The authors concluded that rhBMP-7 is an effective alternative to AICB when supplementing local autograft for 1-level posterolat fusions & avoids the morbidity associated w/ AICB harvesting.	This study population is relatively small (<50 pts), but represents a homogeneous group of pts. The randomization protocol is adequately described & the treating surgeon was blinded as best as possible regarding intervention. Despite randomization there was a difference btwn groups w/ respect to the spinal segment undergoing fusion. Although statistics were adequately described, it is difficult to determine the significance due to the small sample size & no confidence intervals were applied. The small sample size, incomplete statistical analysis, & potential impact of differences in baseline characteristics require that the study be downgraded to Level II evidence in support of rhBMP-7 as a substitute for AICB to supplement local autograft in 1-level instrumented posterolat fusions.
Vaccaro et al., 2008[62]	II	The purpose of the prospective randomized controlled multicenter study was to determine the safety & efficacy of rhBMP-7/ACS & to demonstrate noninferiority as a replacement for AICB for uninstrumented, 1-level posterolat fusion. 335 pts were randomized, but only 293 were treated. 208 pts received rhBMP-7 & 87 control pts received AICB. Independent blinded clinical & radiographic evaluations were performed at 6 wks & at 3, 6, 12, 24, & >36 mos after surgery using validated outcome measures & radiographs. Fusion assessment after 36 mos was supplemented w/ CT scans. At the 24-mo end point the investigational group did not achieve statistical equivalence w/ respect to the overall success rate compared w/ controls (38.7% vs 49.4% [p = 0.33]). The investigational group demonstrated a lower fusion rate (61.7% vs 83.1%). Noninferiority of the overall success was demonstrated after 36 mos (47.2% in the investigational group & 46.8% in the control cohort [p = 0.025]). CT evaluation demonstrated a statistically greater percentage of pts in the control arm demonstrating bridging bone across the intertransverse space (83% vs 56% [p = 0.001]). There was not a statistically significant difference in the rate of treatment-related serious adverse events btwn the investigational & control groups (85.6% & 84.7%, respectively [p = 0.863]). Donor site pain was reported in 44% of controls at 12 mos & 35% after 36 mos, w/ VAS scores of 1.6 at 12 mos & 1.1 after 36 mos. Op time & blood loss were significantly less in the investigational cohort. No significant immunological reaction related to the application of rhBMP-7 was recorded. The authors concluded that rhBMP-7/collagen composite is a safe & effective alternative to ICBG.	This is a well-designed large prospective randomized trial but several limitations exist. Although the authors performed a power calculation to determine sample size, the benefits of randomizations were compromised since 13% of pts were not treated & the authors failed to describe these pts. The follow-up rate at 24 mos was 87% of pts treated (76% of pts randomized), dropping to 69% (60% of randomized pts) after 36 mos. The assessment of overall success was modified after 36 mos compared w/ the 24-mo evaluation. Due to these limitations, this study provides Level II evidence that at 24 mos, rhBMP-7 did not achieve a noninferior outcome to AICB, as measured by the a priori overall success criteria defined by the authors, at 24 mos compared w/ AICB. The results were comparable after 36 mos; however, this conclusion is compromised due to the significant dropout rate.

(continued)

TABLE 6: rhBMP-7 in posterolateral fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Comments
Kanayama et al., 2006	II	The objective of this prospective RCT was to compare the clinical & radiographic outcomes of pts undergoing 1-level posterolat instrumented fusions for degenerative spondylolisthesis. 20 pts were randomized into 2 groups, rhBMP-7/collagen (n = 10) & local autograft/HA/TCP (n = 10). Static & dynamic radiographs & CT images were obtained at 3 & 6 wks as well as at 3, 6, & 12 mos to assess fusion status. Validated clinical outcome measures (ODI) were obtained at 3, 6, 9, & 12 mos after surgery. All pts who demonstrated radiographic fusion underwent removal of instrumentation, regardless of clinical status. The follow-up rate at 1 yr was 90%. Both cohorts demonstrated significant improvement in the ODI scores at 3 mos after surgery; however, no significant difference was observed btwn the groups. It is not clear if the significance of clinical improvement was maintained after this time point. The fusion rate based on radiographic assessment was 78% in the rhBMP-7 cohort & 90% in the control group. Those pts considered fused (7 in the rhBMP-7/collagen cohort & 9 controls) underwent surgical exploration at an average of 15.3 mos after the index procedure. Solid arthrodesis was identified in 57% of rhBMP-7 pts explored (calculated fusion rate for all patients receiving rhBMP-7/collagen = 44%) & in 78% of the control group (calculated fusion rate for entire control group = 70%). No statistical analysis regarding fusion rate was performed. The authors concluded that rhBMP-7 was capable of inducing new bone growth however the fusion rate was not encouraging & they suggested modifications in either surgical technique or carrier may be required.	This is a small pt population, but it benefits from the homogeneous diagnosis. The authors fail to describe the randomization scheme. It is not clear if the radiographic assessment was performed in a blinded fashion. The authors provide cursory baseline demographic data. No power calculation was performed to determine sample size & a superficial description of statistical analysis was performed. The intraop determination of fusion status is generally considered the "gold standard" for fusion assessment & a clear advantage in this study. Due to limitations this study was downgraded to Level II evidence, not supporting the use of rhBMP-7 for 1-level instrumented posterolat fusions in degenerative spondylolisthesis.
Vaccaro et al., 2004	II	The purpose of the prospective randomized controlled multicenter pilot study was to determine the safety & efficacy of rhBMP-7/ACS as a replacement for AICB for uninstrumented posterolat fusion. 36 pts w/ degenerative spondylolisthesis were randomized & followed at 6 wks & 3, 6, & 12 mos after surgery. Safety was determined by comparing the nature & frequency of adverse events. Radiographs were analyzed by independent radiologists, blinded to the intervention, & validated outcomes instruments were used to determine clinical efficacy. The follow-up rate at 12 mos was 79% for the investigational group & 83% for the controls. The rate of adverse events did not differ btwn groups & no specific adverse event was related to the rhBMP-7. The fusion rate for the rhBMP-7 cohort was 74% & for the control group 60% (p = 0.675). The clinical success rate was 86% for the rhBMP-7 cohort & 73% for the control group (p = 0.39). The authors concluded that rhBMP-7 has an acceptable safety profile, & the comparable results to AICB justify further investigation to define the efficacy of rhBMP-7 as a bone graft substitute.	This is a well-designed pilot study but is limited by the small sample size (<50 pts) & pts lost to follow-up at the study end point. No power calculation was performed to determine the sample size. The 1-yr end point may also be considered too brief for fusion procedures. The measure of overall clinical success was a nonvalidated composite score of both fusion & clinical measures. Due to these limitations the study was downgraded to Level II, finding that the use of rhBMP-7 as a bone graft substitute in uninstrumented posterolat fusions is safe & effective.
Vaccaro et al., 2005	II	The purpose of this study was to report the long-term follow-up data from the previously reported pilot study described above. Pts were followed up to 24 mos after surgery using the same clinical & radiographic outcome measures as previously reported. The clinical & radiographic rates of follow-up were 86% & 83%, respectively. The fusion rates for the control & investigational cohort were 40% & 55%, respectively. Clinical success, w/ respect to the validated outcomes instrument, was 85% in the investigational cohort & 64% in the control group. No long-term adverse events were specifically related to the use of rhBMP-7. The authors concluded that safety & efficacy of rhBMP-7 is comparable to AICB for at least 24 mos after surgery.	The initial study (from 2004) was downgraded to Level II evidence due to limitations outlined in the comments. At the 24-mo time period no additional significant limitations were identified; therefore the study maintained a Level II designation supporting rhBMP-7 as a bone graft substitute.

(continued)

TABLE 6: rhBMP-7 in posterolateral fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Comments
Vaccaro et al., 2008[65]	III	The purpose of this study was to report the long-term follow-up data from the previously reported pilot study described above. Pts were followed up to 48 mos after surgery using the same clinical & radiographic outcome measures as previously reported. The clinical & radiographic rates of follow-up were 69% & 61%, respectively. The fusion rates for the control & investigational cohort were 50% & 68.8%, respectively. Clinical success, w/ respect to the validated outcomes instrument, was 73.7% in the investigational cohort & 57.1% in the control group. The overall success was 62.5% in the investigational cohort & 33.3% in the control group. No long-term adverse events were specifically related to the use of rhBMP-7. The authors concluded that safety & efficacy of rhBMP-7 is comparable to AICB for at least 48 mos after surgery.	The initial study (from 2004) was downgraded to Level II evidence due to limitations outlined in the comments. Further limitations were identified during the follow-up period that significantly limit conclusions drawn from this study, in particular the large no. of pts lost to follow-up. Due to this significant limitation this follow-up report provides only Level III evidence supporting rhBMP-7 as a bone graft substitute.
Vaccaro et al., 2003	IV	The purpose of this pilot study was to determine the safety of rhBMP-7 combined w/ autograft for posterolat uninstrumented fusions in pts w/ symptomatic degenerative spondylolisthesis. 12 pts underwent a 1-level fusion receiving rhBMP-7 mixed w/ AICB. Independent radiographic assessment of dynamic radiographs was performed at 6 wks & 3, 6, 9, & 12 mos, & validated outcomes instruments, ODI scores, were used to assess clinical outcome at 6 & 12 mos. Results were compared w/ a historical cohort of pts who only received AICB. 75% of pts achieved clinical success & 55% attained a solid fusion. There were no adverse events specifically associated w/ the use of rhBMP-7. The authors concluded that rhBMP-7 demonstrated an acceptable safety profile when used as an adjunct to AICB.	The small pt population, compared w/ a historical cohort, & relatively short follow-up for a fusion procedure limit this cohort study. The study design is difficult to categorize but falls btwn a prospective case series & retrospective cohort study. It was initially considered a Level III study but due to the limitations was downgraded to Level IV evidence supporting the safety & efficacy of rhBMP-7 as a bone graft extender for uninstrumented 1-level posterolat fusions in pts w/ symptomatic spondylolisthesis.

* ACS = absorbable collagen sponge; AICB = autologous iliac crest bone; HA = hydroxyapatite; ICBG = iliac crest bone graft; ODI = Oswestry Disability Index; pt = patient; RCT = randomized controlled trial; rhBMP-7 = recombinant human bone morphogenetic protein–7; TCP = tricalcium phosphate; VAS = visual analog scale.

lower fusion rate (61.7% vs 83.1%). Noninferiority of the overall success was demonstrated after 36 months (47.2% in the investigational group and 46.8% in the control cohort [p = 0.025]). Computed tomography evaluation demonstrated a statistically greater percentage of patients in the control arm demonstrating bridging bone across the intertransverse space (83% vs 56% [p = 0.001]). However, at 36 months after surgery, only 69% of the original 293 patients were available for evaluation. There was no significant difference in the rate of treatment-related serious adverse events between the investigational and control groups, 85.6% and 84.7% respectively (p = 0.863). No significant immunological reaction related to the application of rhBMP-7 was recorded. The authors concluded that rhBMP-7/collagen composite is a safe and effective alternative to ICBG. This was a well-designed and executed trial; however, limitations exist. The randomization process was compromised as 13% of those originally randomized were not treated, and no data are provided regarding these patients or those lost at final follow-up. The method of fusion assessment was altered at the final follow-up, with the addition of CT images, and a significant number of patients not available for evaluation (> 30%). Due to these limitations the study was downgraded to Level II evidence that rhBMP-7/ACS is noninferior to AICB for noninstrumented posterolateral lumbar fusion.

Kanayama et al. conducted a prospective RCT to compare the clinical and radiographic outcomes in 20 patients with degenerative spondylolisthesis receiving either rhBMP-7/ACS (n = 10) or local autograft/HA/TCP (n = 10) for single-level posterolateral instrumented fusions.[35] Fusion status was assessed using plain radiographs and CT images at 3 and 6 weeks as well as at 3, 6, and 12 months after surgery. Validated clinical outcome measures (ODI) were obtained at 3, 6, 9, and 12 months after surgery. Regardless of clinical status, all patients underwent a second surgery to remove their instrumentation if a solid arthrodesis was diagnosed based on radiographic imaging, on average 15.3 months following the index procedure. At 1 year after surgery the follow-up rate was 90%. The ODI scores significantly improved at 3 months in both cohorts; however, it is difficult to determine if the significance of this improvement was maintained beyond 3 months. The fusion rate based on radiographic assessment was 78% in the rhBMP-7 cohort; however, only 57% of these patients demonstrated a solid arthrodesis during direct surgical exploration. The radiographic fusion rate of the control group was 90%, with 78% of controls demonstrating a fusion at the time of implant removal. No statistical analysis regarding fusion rate was performed. The authors concluded that utilization of rhBMP-7 was feasible, but the observed fusion rate was not encouraging, suggesting that modifications of the surgical technique or carrier were required. The study cohort was small yet homogeneous with respect to surgical procedure and presenting diagnosis. It is not clear if the radiographic assessment was performed in a blinded fashion; however, confirmation of fusion through direct operative exploration is considered the gold standard for fusion assessment. Limitations of the study design include failure to describe the randomization scheme or perform a power calculation to determine sample size. The authors failed to perform an adequate statistical analysis; this may be a secondary consequence of the small sample size. Despite the randomized nature of this study, the study was downgraded to Level II evidence suggesting that rhBMP-7 is an inadequate substitute for AICB in instrumented posterolateral fusion.

Vaccaro et al. published 3 separate studies over a 4-year period reporting the radiographic and clinical results from a prospective randomized controlled multicenter clinical pilot study investigating the efficacy and safety of OP-1 compared with AICB in noninstrumented posterolateral lumbar fusions.[61,64,65] The original pilot study, published in 2004, randomized 36 patients with degenerative spondylolisthesis to receive either rhBMP-7/ACS or AICB.[64] The initial study followed patients at 6 weeks and at 3, 6, and 12 months after the index procedure. An independent blinded radiologist evaluated plain radiographs to assess fusion and validated outcome measures; ODI and SF-36, were used to assess clinical status. At 12 months after surgery, the follow-up rate was 79% for the rhBMP-7 cohort and 83% for the control group. No short-term adverse events directly related to the use of rhBMP-7 were reported. At 12 months after surgery, the fusion and clinical success rates were 74% and 86%, respectively, for the rhBMP-7 cohort and 60% and 73%, respectively, in the control group, with no statistically significant difference between groups. From this initial study the authors concluded that rhBMP-7 has an acceptable safety profile and comparable results to AICB to justify further investigation. This initial study was limited by the lack of a power calculation, a small sample size (< 50 patients), and significant loss to follow-up within the interventional group. A nonvalidated overall clinical outcome composite score is included that is difficult to objectify. Given these limitations, and those inherent with a pilot study, the initial publication is downgraded to Level II evidence in support of an acceptable safety profile and comparable efficacy of rhBMP-7 to AICB.

The same authors published 2 follow-up studies in 2005 and 2008 to report the 2- and 4-year outcomes from the same study population.[61,65] As the initial report provided Level II evidence, the subsequent reports were started at this level and downgraded further if additional limitations were identified. At the 24-month follow-up end point, the follow-up rate for the investigational and control groups were 86% and 83%, respectively.[61] Radiographic fusion occurred in 55% of patients receiving OP-1 and 40% of AICB patients. Clinical success was recorded in 85% of OP-1 patients and 64% of control patients. No additional limitations were identified in this study; therefore, a Level II designation is maintained. However, at 48 months only 69% and 61% were available for clinical and radiographic evaluations, respectively.[65] The fusion and clinical success rates were reported as 68.8% and 73.7% in the rh-BMP-7 group and 50% and 57.1% in the control cohort. Due to the small number of patients available at 48 months, formal statistical analysis was not performed. The authors concluded that rhBMP-7 had an acceptable safety profile and comparable results to AICB. Due to the additional attrition of patients at the 48-month follow-up time point, this study was downgraded to Level III evidence in support of comparable efficacy between rhBMP-7 and AICB.

In 2003, Vaccaro et al. conducted a pilot study to determine the safety of rhBMP-7 combined with AICB for posterolateral uninstrumented fusions in patients with symptomatic degenerative spondylolisthesis.[63] Seventy-five percent of patients achieved clinical success and 55% attained a solid fusion. There were no adverse events specifically associated with the use of rhBMP-7. The authors concluded that rhBMP-7 demonstrated an acceptable safety profile when used as an adjunct to AICB. The small patient population, comparison with a historical cohort, and relatively short follow-up for a fusion procedure limit the study. Due to these limitations the study is downgraded to Level IV evidence. As this was the only study investigating this specific application of rhBMP-7, there is insufficient evidence to formulate a recommendation.

Summary

A wide variety of bone graft extenders and substitutes are currently available. Enhanced fusion rates and the ability to avoid complications associated with iliac crest harvesting are the intended benefits of their use. Many, if not all, of these extenders and substitutes evaluated in this review have demonstrated a positive effect on fusion rate with clinical outcomes comparable to AICB. Convincing evidence exists that calcium-based composites cannot be considered substitutes for AICB due to inferior fusion rates. These materials, along with allograft-derived grafts (DBM), function primarily as extenders, requiring some form of autograft to achieve adequate fusion rates. There has been little if any risk associated with the use of these extenders.

Bone morphogenetic proteins have dramatically altered the landscape of spinal fusion surgery. These powerful osteoinductive agents have demonstrated excellent potential as substitutes for AICB with both interbody and posterolateral fusions. The vast majority of investigations have evaluated the effect of rhBMP-2. Although rhBMP-2 has shown a positive effect on fusion rate, complications have been reported related to its use. As a result, careful consideration is required when utilizing these products.

Despite the beneficial effect on fusion, the current literature has also not adequately addressed the issue of whether these improved fusion rates justify the cost, especially for treatment of routine degenerative lumbar disease. Although it is likely that certain patient populations would benefit from the addition of BMPs when performing spinal fusion surgery, the current literature has failed to adequately identify such patient populations.

Key Issues for Future Investigation

There has already been an extensive amount of research investigating the potential impact of these graft extenders and substitutes. Further investigations should focus on improving study design to validate the conclusions formulated from previous publications and a comprehensive evaluation of risks and complications will be necessary to properly inform our patients. Identification of patient populations at risk for pseudarthrosis would also better define patient populations where the benefits of utilizing BMPs justify the risks. Potentially more relevant than defining the clinical impact of these materials is to determine their cost utility. Comprehensive cost analyses, not simply a superficial quantification of upfront costs, will ultimately be required. Such an endeavor will require the concerted effort of a multidisciplinary panel of experts from the clinical, epidemiological, and administrative disciplines.

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., 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 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: Kaiser. 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: Kaiser. Study supervision: Kaiser.

References

1. Acharya NK, Kumar RJ, Varma HK, Menon VK: Hydroxyapatite-bioactive glass ceramic composite as stand-alone graft substitute for posterolateral fusion of lumbar spine: a prospective, matched, and controlled study. **J Spinal Disord Tech 21:**106–111, 2008
2. Agarwal R, Williams K, Umscheid CA, Welch WC: Osteoinductive bone graft substitutes for lumbar fusion: a systematic review. Clinical article. **J Neurosurg Spine 11:**729–740, 2009
3. Balseiro S, Nottmeier EW: Vertebral osteolysis originating from subchondral cyst end plate defects in transforaminal lumbar interbody fusion using rhBMP-2. Report of two cases. **Spine J 10:**e6–e10, 2010
4. Banwart JC, Asher MA, Hassanein RS: Iliac crest bone graft harvest donor site morbidity. A statistical evaluation. **Spine (Phila Pa 1976) 20:**1055–1060, 1995

5. 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
6. Burkus JK, Gornet MF, Dickman CA, Zdeblick TA: Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. **J Spinal Disord Tech 15:**337–349, 2002
7. Burkus JK, Gornet MF, Schuler TC, Kleeman TJ, Zdeblick TA: Six-year outcomes of anterior lumbar interbody arthrodesis with use of interbody fusion cages and recombinant human bone morphogenetic protein-2. **J Bone Joint Surg Am 91:**1181–1189, 2009
8. Cammisa FP Jr, Lowery G, Garfin SR, Geisler FH, Klara PM, McGuire RA, et al: Two-year fusion rate equivalency between Grafton DBM gel and autograft in posterolateral spine fusion: a prospective controlled trial employing a side-by-side comparison in the same patient. **Spine (Phila Pa 1976) 29:**660–666, 2004
9. Carragee EJ, Ghanayem AJ, Weiner BK, Rothman DJ, Bono CM: A challenge to integrity in spine publications: years of living dangerously with the promotion of bone growth factors. **Spine J 11:**463–468, 2011
10. Carragee EJ, Mitsunaga KA, Hurwitz EL, Scuderi GJ: Retrograde ejaculation after anterior lumbar interbody fusion using rhBMP-2: a cohort controlled study. **Spine J 11:**511–516, 2011
11. Carter JD, Swearingen AB, Chaput CD, Rahm MD: Clinical and radiographic assessment of transforaminal lumbar interbody fusion using HEALOS collagen-hydroxyapatite sponge with autologous bone marrow aspirate. **Spine J 9:**434–438, 2009
12. Chang CH, Lin MZ, Chen YJ, Hsu HC, Chen HT: Local autogenous bone mixed with bone expander: an optimal option of bone graft in single-segment posterolateral lumbar fusion. **Surg Neurol 70 Suppl 1:**S47–S49, 2008
13. Chen CL, Liu CL, Sun SS, Han PY, Lee CS, Lo WH: Posterolateral lumbar spinal fusion with autogenous bone chips from laminectomy extended with OsteoSet. **J Chin Med Assoc 69:**581–584, 2006
14. Chen NF, Smith ZA, Stiner E, Armin S, Sheikh H, Khoo LT: Symptomatic ectopic bone formation after off-label use of recombinant human bone morphogenetic protein-2 in transforaminal lumbar interbody fusion. Report of 4 cases. **J Neurosurg Spine 12:**40–46, 2010
15. Chen WJ, Tsai TT, Chen LH, Niu CC, Lai PL, Fu TS, et al: The fusion rate of calcium sulfate with local autograft bone compared with autologous iliac bone graft for instrumented short-segment spinal fusion. **Spine (Phila Pa 1976) 30:**2293–2297, 2005
16. Dai LY, Jiang LS: Single-level instrumented posterolateral fusion of lumbar spine with beta-tricalcium phosphate versus autograft: a prospective, randomized study with 3-year follow-up. **Spine (Phila Pa 1976) 33:**1299–1304, 2008
17. Dawson E, Bae HW, Burkus JK, Stambough JL, Glassman SD: Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. **J Bone Joint Surg Am 91:**1604–1613, 2009
18. Delawi D, Dhert WJ, Rillardon L, Gay E, Prestamburgo D, Garcia-Fernandez C, et al: A prospective, randomized, controlled, multicenter study of osteogenic protein-1 in instrumented posterolateral fusions: report on safety and feasibility. **Spine (Phila Pa 1976) 35:**1185–1191, 2010
19. Dimar JR, Glassman SD, Burkus KJ, Carreon LY: Clinical outcomes and fusion success at 2 years of single-level instrumented posterolateral fusions with recombinant human bone morphogenetic protein-2/compression resistant matrix versus iliac crest bone graft. **Spine (Phila Pa 1976) 31:**2534–2540, 2006
20. Dimar JR II, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY: Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. **J Bone Joint Surg Am 91:**1377–1386, 2009
21. Epstein NE: Beta tricalcium phosphate: observation of use in 100 posterolateral lumbar instrumented fusions. **Spine J 9:**630–638, 2009
22. Epstein NE: Efficacy of different bone volume expanders for augmenting lumbar fusions. **Surg Neurol 69:**16–19, 2008
23. Garrett MP, Kakarla UK, Porter RW, Sonntag VKH: Formation of painful seroma and edema after the use of recombinant human bone morphogenetic protein-2 in posterolateral lumbar spine fusions. **Neurosurgery 66:**1044–1049, 2010
24. Geibel PT, Boyd DL, Slabisak V: The use of recombinant human bone morphogenic protein in posterior interbody fusions of the lumbar spine: a clinical series. **J Spinal Disord Tech 22:**315–320, 2009
25. Glassman SD, Carreon L, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, et al: Posterolateral lumbar spine fusion with INFUSE bone graft. **Spine J 7:**44–49, 2007
26. Glassman SD, Carreon LY, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, et al: RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. **Spine (Phila Pa 1976) 33:**2843–2849, 2008
27. Glassman SD, Dimar JR III, Burkus K, Hardacker JW, Pryor PW, Boden SD, et al: The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers. **Spine (Phila Pa 1976) 32:**1693–1698, 2007
28. Goldberg VM, Stevenson S: Natural history of autografts and allografts. **Clin Orthop Relat Res (225):**7–16, 1987
29. Haid RW Jr, Branch CL Jr, Alexander JT, Burkus JK: Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. **Spine J 4:**527–539, 2004
30. Hamilton DK, Jones-Quaidoo SM, Sansur C, Shaffrey CI, Oskouian R, Jane JA Sr: Outcomes of bone morphogenetic protein-2 in mature adults: posterolateral non-instrument-assisted lumbar decompression and fusion. **Surg Neurol 69:**457–462, 2008
31. Heary RF, Schlenk RP, Sacchieri TA, Barone D, Brotea C: Persistent iliac crest donor site pain: independent outcome assessment. **Neurosurgery 50:**510–517, 2002
32. Hsu CJ, Chou WY, Teng HP, Chang WN, Chou YJ: Coralline hydroxyapatite and laminectomy-derived bone as adjuvant graft material for lumbar posterolateral fusion. **J Neurosurg Spine 3:**271–275, 2005
33. Jenis LG, Banco RJ: Efficacy of silicate-substituted calcium phosphate ceramic in posterolateral instrumented lumbar fusion. **Spine (Phila Pa 1976) 35:**E1058–E1063, 2010
34. Joseph V, Rampersaud YR: Heterotopic bone formation with the use of rhBMP2 in posterior minimal access interbody fusion: a CT analysis. **Spine (Phila Pa 1976) 32:**2885–2890, 2007
35. Kanayama M, Hashimoto T, Shigenobu K, Yamane S, Bauer TW, Togawa D: A prospective randomized study of posterolateral lumbar fusion using osteogenic protein-1 (OP-1) versus local autograft with ceramic bone substitute: emphasis of surgical exploration and histologic assessment. **Spine (Phila Pa 1976) 31:**1067–1074, 2006
36. Katayama Y, Matsuyama Y, Yoshihara H, Sakai Y, Nakamura H, Imagama S, et al: Clinical and radiographic outcomes of posterolateral lumbar spine fusion in humans using recombinant human bone morphogenetic protein-2: an average five-year follow-up study. **Int Orthop 33:**1061–1067, 2009
37. Knox JB, Dai JM III, Orchowski J: Osteolysis in transforaminal lumbar interbody fusion with bone morphogenetic protein-2. **Spine (Phila Pa 1976) 36:**672–676, 2011

38. Korovessis P, Koureas G, Zacharatos S, Papazisis Z, Lambiris E: Correlative radiological, self-assessment and clinical analysis of evolution in instrumented dorsal and lateral fusion for degenerative lumbar spine disease. Autograft versus coralline hydroxyapatite. **Eur Spine J 14:**630–638, 2005
39. Kurz LT, Garfin SR, Booth RE Jr: Harvesting autogenous iliac bone grafts. A review of complications and techniques. **Spine (Phila Pa 1976) 14:**1324–1331, 1989
40. Lanman TH, Hopkins TJ: Lumbar interbody fusion after treatment with recombinant human bone morphogenetic protein-2 added to poly(L-lactide-co-D,L-lactide) bioresorbable implants. **Neurosurg Focus 16(3):**E9, 2004
41. Lee JH, Hwang CJ, Song BW, Koo KH, Chang BS, Lee CK: A prospective consecutive study of instrumented posterolateral lumbar fusion using synthetic hydroxyapatite (Bongros-HA) as a bone graft extender. **J Biomed Mater Res A 90:**804–810, 2009
42. Lewandrowski KU, Nanson C, Calderon R: Vertebral osteolysis after posterior interbody lumbar fusion with recombinant human bone morphogenetic protein 2: a report of five cases. **Spine J 7:**609–614, 2007
43. Mindea SA, Shih P, Song JK: Recombinant human bone morphogenetic protein-2-induced radiculitis in elective minimally invasive transforaminal lumbar interbody fusions: a series review. **Spine (Phila Pa 1976) 34:**1480–1485, 2009
44. Mummaneni PV, Pan J, Haid RW, Rodts GE: Contribution of recombinant human bone morphogenetic protein-2 to the rapid creation of interbody fusion when used in transforaminal lumbar interbody fusion: a preliminary report. **J Neurosurg Spine 1:**19–23, 2004
45. Neen D, Noyes D, Shaw M, Gwilym S, Fairlie N, Birch N: Healos and bone marrow aspirate used for lumbar spine fusion: a case controlled study comparing healos with autograft. **Spine (Phila Pa 1976) 31:**E636–E640, 2006
46. Ong KL, Villarraga ML, Lau E, Carreon LY, Kurtz SM, Glassman SD: Off-label use of bone morphogenetic proteins in the United States using administrative data. **Spine (Phila Pa 1976) 35:**1794–1800, 2010
47. Pradhan BB, Bae HW, Dawson EG, Patel VV, Delamarter RB: Graft resorption with the use of bone morphogenetic protein: lessons from anterior lumbar interbody fusion using femoral ring allografts and recombinant human bone morphogenetic protein-2. **Spine (Phila Pa 1976) 31:**E277–E284, 2006
48. 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 16: bone graft extenders and substitutes. **J Neurosurg Spine 2:** 733–736, 2005
49. Rihn JA, Makda J, Hong J, Patel R, Hilibrand AS, Anderson DG, et al: The use of RhBMP-2 in single-level transforaminal lumbar interbody fusion: a clinical and radiographic analysis. **Eur Spine J 18:**1629–1636, 2009
50. Rihn JA, Patel R, Makda J, Hong J, Anderson DG, Vaccaro AR, et al: Complications associated with single-level transforaminal lumbar interbody fusion. **Spine J 9:**623–629, 2009
51. Schizas C, Triantafyllopoulos D, Kosmopoulos V, Tzinieris N, Stafylas K: Posterolateral lumbar spine fusion using a novel demineralized bone matrix: a controlled case pilot study. **Arch Orthop Trauma Surg 128:**621–625, 2008
52. Shah RK, Moncayo VM, Smitson RD, Pierre-Jerome C, Terk MR: Recombinant human bone morphogenetic protein 2-induced heterotopic ossification of the retroperitoneum, psoas muscle, pelvis and abdominal wall following lumbar spinal fusion. **Skeletal Radiol 39:**501–504, 2010
53. Singh K, Smucker JD, Gill S, Boden SD: Use of recombinant human bone morphogenetic protein-2 as an adjunct in posterolateral lumbar spine fusion: a prospective CT-scan analysis at one and two years. **J Spinal Disord Tech 19:**416–423, 2006 (Erratum in **J Spinal Disord Tech 20:**185, 2007)
54. Slosar PJ, Josey R, Reynolds J: Accelerating lumbar fusions by combining rhBMP-2 with allograft bone: a prospective analysis of interbody fusion rates and clinical outcomes. **Spine J 7:** 301–307, 2007
55. Summers BN, Eisenstein SM: Donor site pain from the ilium. A complication of lumbar spine fusion. **J Bone Joint Surg Br 71:**677–680, 1989
56. Taghavi CE, Lee KB, Keorochana G, Tzeng ST, Yoo JH, Wang JC: Bone morphogenetic protein-2 and bone marrow aspirate with allograft as alternatives to autograft in instrumented revision posterolateral lumbar spinal fusion: a minimum two-year follow-up study. **Spine (Phila Pa 1976) 35:**1144–1150, 2010
57. Urist MR: Bone: formation by autoinduction. **Science 150:**893–899, 1965
58. US Food and Drug Administration: **InFUSE™ Bone Graft/LT-CAGE™ Lumbar Tapered Fusion Device - P000058.** http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/DeviceApprovalsandClearances/Recently-ApprovedDevices/ucm083423.htm) [Accessed April 21, 2014]
59. US Food and Drug Administration: **Listing of CDRH Humanitarian Device Exemptions.** (http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/DeviceApprovalsandClearances/HDEApprovals/ucm161827.htm) [Accessed April 21, 2014]
60. US Food and Drug Administration: **OP-1 Putty - H0200008.** (http://www.fda.gov/MedicalDevices/ProductsandMedicalrocedures/DeviceApprovalsandClearances/Recently-ApprovedDevices/ucm081181.htm) [Accessed April 21, 2014]
61. Vaccaro AR, Anderson DG, Patel T, Fischgrund J, Truumees E, Herkowitz HN, et al: Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. **Spine (Phila Pa 1976) 30:**2709–2716, 2005
62. Vaccaro AR, Lawrence JP, Patel T, Katz LD, Anderson DG, Fischgrund JS, et al: The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis: a long-term (>4 years) pivotal study. **Spine (Phila Pa 1976) 33:**2850–2862, 2008
63. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz H, et al: A pilot safety and efficacy study of OP-1 putty (rhBMP-7) as an adjunct to iliac crest autograft in posterolateral lumbar fusions. **Eur Spine J 12:**495–500, 2003
64. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz HN, et al: A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. **Spine (Phila Pa 1976) 29:**1885–1892, 2004
65. Vaccaro AR, Whang PG, Patel T, Phillips FM, Anderson DG, Albert TJ, et al: The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft for posterolateral lumbar arthrodesis: minimum 4-year follow-up of a pilot study. **Spine J 8:**457–465, 2008
66. Vaidya R, Weir R, Sethi A, Meisterling S, Hakeos W, Wybo CD: Interbody fusion with allograft and rhBMP-2 leads to consistent fusion but early subsidence. **J Bone Joint Surg Br 89:**342–345, 2007
67. Villavicencio AT, Burneikiene S, Nelson EL, Bulsara KR, Favors M, Thramann J: Safety of transforaminal lumbar interbody fusion and intervertebral recombinant human bone morphogenetic protein-2. **J Neurosurg Spine 3:**436–443, 2005

Manuscript submitted March 27, 2014.
Accepted April 9, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14325.

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.

J Neurosurg Spine 21:133–139, 2014

Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 17: Bone growth stimulators as an adjunct for lumbar fusion

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

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

The relationship between the formation of a solid arthrodesis and electrical and electromagnetic energy is well established; most of the information on the topic, however, pertains to the healing of long bone fractures. The use of both invasive and noninvasive means to supply this energy and supplement spinal fusions has been investigated. Three forms of electrical stimulation are routinely used: direct current stimulation (DCS), pulsed electromagnetic field stimulation (PEMFS), and capacitive coupled electrical stimulation (CCES). Only DCS requires the placement of electrodes within the fusion substrate and is inserted at the time of surgery. Since publication of the original guidelines, few studies have investigated the use of bone growth stimulators. Based on the current review, no conflict with the previous recommendations was generated. The use of DCS is recommended as an option for patients younger than 60 years of age, since a positive effect on fusion has been observed. The same, however, cannot be stated for patients over 60, because DCS did not appear to have an impact on fusion rates in this population. No study was reviewed that investigated the use of CCES or the routine use of PEMFS. A single low-level study demonstrated a positive impact of PEMFS on patients undergoing revision surgery for pseudarthrosis, but this single study is insufficient to recommend for or against the use of PEMFS in this patient population.
(http://thejns.org/doi/abs/10.3171/2014.4.SPINE14326)

Key Words • **lumbar spine** • **lumbar fusion** • **bone growth stimulator** • **practice guidelines**

Recommendations

There is no evidence that conflicts with the previous recommendations regarding bone growth stimulation published in the original version of the "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine."[18]

Grade C

The routine use of DCS in patients over the age of 60 years is not recommended, as the evidence demonstrates no impact on fusion rates (single Level II study).

For patients younger than 60 years of age, undergoing a lumbar fusion, the use of DCS is an option as studies have demonstrated a positive impact on fusion rate; however, there is insufficient evidence regarding its impact on clinical outcome (single Level III study/multiple Level IV studies).

Grade I

There is insufficient evidence to recommend for or against the use of PEMFS as a treatment alternative to re-

Abbreviations used in this paper: BMP = bone morphogenetic protein; CCES = capacitive coupled electrical stimulation; DCS = direct current stimulation; DEXA = dual energy x-ray absorptiometry; DPQ = Dallas Pain Questionnaire; LBPRS = Low Back Pain Rating Scale; PEMFS = pulsed electromagnetic field stimulation; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale.

vision surgery in patients presenting with pseudarthrosis following posterior lumbar fusion (single Level IV study).

Rationale

Since the publication of the original "Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine,"[16] the evidence supporting the role of lumbar fusion as an effective treatment alternative for a variety of degenerative spinal conditions continues to expand.[6,9,12,22] As the role of lumbar fusion becomes more established, increasing emphasis has been placed on maneuvers to enhance the potential for a solid arthrodesis. The positive impact of spinal instrumentation on fusion rates is well recognized.[5,11,17] There is also a growing body of evidence demonstrating a beneficial effect on fusion rates with osteoinductive agents.[1,7,8] The data supporting the role of bone growth stimulators remain inconclusive and more controversial.[15,18]

The interaction between electrical energy and the formation of an osseous union is a well-recognized concept, with the majority of clinical data focusing on long bone healing.[15] Dwyer published one of the first manuscripts describing the utilization of direct current stimulation (DCS) for spinal fusion.[10] Since this report, 3 forms of electrical stimulation have gained acceptance for use in spinal fusion: DCS, pulsed electromagnetic field stimulation (PEMFS), and capacitive coupled electrical stimulation (CCES). DCS requires the insertion of cathodes, attached to an implanted battery, directly into the fusion substrate. PEMFS is a noninvasive means of delivering electromagnetic energy to the fusion by wearing an external coil driven by an electrical current. CCES relies on the generation of an electrical field through capacitive plates placed on the patient's skin.[15] The purpose of this update was to review the current literature and examine the evidence supporting the clinical utility of various bone growth stimulators for lumbar fusion surgery, although no studies investigating the efficacy of CCES were identified.

Search Criteria

A computerized search of the National Library of Medicine MEDLINE database, utilizing the online search engine PubMed, was conducted for the period from 2003 through December 2011 utilizing the following search terms ((("Lumbosacral Region"[MeSH] OR "Lumbar Vertebrae"[MeSH]) AND "Spinal Fusion"[MeSH]) OR "lumbar fusion"[All Fields] OR ("lumbar"[title] AND "fusion"[title])) AND ((bone growth stimulator[title] OR bone growth stimulators[title]) OR ("Electric Stimulation"[MeSH] OR "Electric Stimulation Therapy"[MeSH] OR (("bone and bones"[MeSH] OR ("bone"[All Fields] AND "bones"[All Fields]) OR "bone and bones"[All Fields] OR "bone"[All Fields]) AND stimulator[All Fields]) OR (("bone and bones"[MeSH] OR ("bone"[All Fields] AND "bones"[All Fields]) OR "bone and bones"[All Fields] OR "bone"[All Fields]) AND stimulators[All Fields]))). The search was limited to the English language and human subjects and yielded a total of 44 articles. The titles and abstracts of these publications were reviewed and those specifically investigating the clinical efficacy of bone growth stimulation were selected. A secondary review of the bibliographies of these articles was conducted to identify any additional relevant manuscripts. A total of 5 manuscripts were selected and serve as the scientific foundation for the updated review.

Scientific Foundation

Andersen et al. performed a randomized, controlled, multicenter trial to determine the impact of DCS on functional outcome of noninstrumented lumbar fusion for patients over 60 years of age.[3] One hundred seven patients presenting with a variety of spinal degenerative disorders and undergoing single or multilevel posterolateral lumbar fusion (PLF) with local autograft and allograft were randomized into cohorts with a 40-mA (n = 44) or 100-mA (n = 11) DCS implanted stimulator or without (n = 43) DCS. For a variety of reasons, 9 randomized patients were excluded either prior to surgery or due to intraoperative complications. Patients completed a series of validated, objective outcome instruments (the 36-Item Short Form Health Survey [SF-36], the Dallas Pain Questionnaire [DPQ], and the Low Back Pain Rating Scale [LBPRS]), and statistical analysis was performed to compare treatment effect. Patients were followed up for 2 years; however, 27% of patients did not complete the functional outcome questionnaires at this end point. At the 2-year point, the patients in the combined treatment group demonstrated significantly greater improvement in 3 of the 4 domains of the DPQ, although no significant difference in LBPRS or SF-36 scores was observed. Based on these results the authors concluded that surgery led to an improvement in functional outcome and that DCS may have a beneficial effect on lumbar fusion in older patients. This is a relatively well-designed randomized trial, but the study does suffer from several limitations. The validity of separating the results of the individual domains within the DPQ is unclear because the overall percentages are graphed to create a profile summary of the patient.[14] Variability existed with respect to the presenting diagnosis and surgical intervention. It is not clear who performed the functional assessment and whether that individual was blinded to the treatment received. At the 2-year follow-up, only 73% of the participants completed the functional assessment questionnaires. Finally, the statistical analysis was limited by the authors' failure to determine the confidence intervals for the observed results. Due to these limitations the study was downgraded to Level II evidence supporting the role of DCS for this patient population undergoing noninstrumented lumbar fusion (Table 1).

Anderson and colleagues published 2 additional studies based on the same patient population with the intention of determining the effect of DCS on fusion rate, correlating the radiographic outcome to clinical outcome, and clarifying whether DCS had an impact on the quality of fusion.[2,4] Of the original 107 patients randomized, 95 were available for fusion assessment at 1 year and 84 were available at 2 years. Thin-slice CT images and plain radiographs were used to assess fusion status. In both the control and treatment cohorts the observed fusion rate was

TABLE 1: Bone growth stimulators as an adjunct for lumbar fusion: summary of evidence*

Authors & Year	Level of Evidence	Description	Comments
Andersen et al., 2009[3]	II	This is a randomized, controlled, multicenter trial to investigate the impact of DCS on functional outcome in pts over the age of 60 yrs undergoing noninstrumented lumbar fusion. 107 pts presenting w/ various spinal degenerative disorders & undergoing PLF w/ local autograft & allograft were randomized into cohorts w/o (n = 43) or w/ the insertion of a 40-mA (n = 44) or 10-mA (n = 11) DC electrical stimulator. Validated, objective outcome instruments (SF-36, DPQ, LBPRS) were utilized to evaluate the functional outcome, & statistical analysis was performed to compare treatment effect. Patients were followed for 2 yrs, but 27% did not complete the functional outcome questionnaires at this time point. At 2 yrs' follow-up the pts in the combined treatment group demonstrated significantly greater improvement in 3 of the 4 domains of the DPQ; however, no significant difference in LBPRS or SF-36 was observed. The authors concluded that surgery led to an improvement in functional outcome & that DCS may have a positive effect on fusions in this pt population.	This is a relatively well-designed randomized trial. There was some variability regarding surgical intervention. Participant flow & care providers are incompletely described. It is not clear who tabulated the functional outcome measures & whether the assessor was blinded to the intervention. Only 73% of pts were available for 2-yr follow-up. The authors failed to provide confidence intervals w/ the statistical analysis.
Andersen et al., 2009[2]	II	This companion study to the authors' original investigation was intended to determine the effect of DCS on fusion status & correlate radiographic outcome w/ clinical outcome. From the original 107 pts randomized, 95 were available for fusion assessment at 1 yr, & 84 were available at 2 yrs. Fusion status was evaluated through thin-slice CT & plain radiographs. Fusion rates were low in the control & treatment cohorts (33% & 32%, respectively). The insertion of a DC stimulator had no impact on fusion rate. Functional outcome correlated w/ presence of a solid arthrodesis. There was a poor correlation of fusion assessment btwn CT & plain radiographs. The authors concluded that DCS had no significant impact on fusion rate; however, a solid fusion defined by CT resulted in better functional outcome & less pain.	The study's limitations are highlighted above. This arm of the study benefited from the blinded radiographic evaluation of the imaging studies. The authors claim that 89% of pts were available at 2-yr follow-up, but this percentage was calculated based on the 95 pts who underwent imaging at 1 yr rather than on the original 107 pts randomized into study cohorts. No information is available regarding the participants not available for 1-yr follow-up.
Rogozinski et al., 2009	III	This single-center, prospective, nonrandomized trial compared radiographic outcome of a heterogeneous group of 31 pts undergoing instrumented PLF supplemented w/ either BMP or an implanted spinal fusion stimulator. Fusion status was assessed using plain radiographs &/or CT, & pain status was determined through a 10-point VAS. The BMP cohort demonstrated a 100% fusion rate while the stimulator group demonstrated a 93.4% fusion rate. BMP cohort was considered to achieve more robust fusion & at a faster rate than the stimulator cohort. Pain improved in both cohorts. The authors concluded that use of BMP led to more rapid graft maturation & more robust fusion compared to fusions supplemented w/ an internal stimulator.	This is a poorly designed & conducted cohort study consisting of an exceedingly small heterogeneous population of pts. Inclusion & exclusion criteria were not defined. A variety of surgical procedures were implemented. The no. of pts available for follow-up is not provided; however, only 9 pts were available for CT analysis at the 2-yr study end point. This would indicate a loss to follow-up of 79%. Radiographic evaluation was not blinded but performed by the operating surgeons, & the criteria for fusion were not defined. Statistical analysis was not adequately described & no confidence intervals were provided.
Simmons et al., 2004	IV	This case series involved pts who presented w/ pseudarthrosis after attempted lumbar fusion & were treated w/ PEMFS. 25 investigators from multiple institutions enrolled 100 pts who received PEMFS for at least 90 days. Radiographic evaluation of fusion was performed by the investigators as well as a blinded radiologist and, if disagreement occurred, a blinded orthopedist. Fusion success rate was 67%, w/ 63% of these pts demonstrating an excellent or good outcome. Only 30% of pts w/ persistent pseudarthrosis demonstrated an excellent or good outcome. The authors concluded that PEMFS was an effective alternative to revision surgery for pts presenting w/ pseudarthrosis.	This case series provides information that demonstrates feasibility of PEMFS for the treatment of pseudarthrosis; however, the true treatment effect cannot be determined given the lack of an appropriate control group. Any comparison to alternative therapies is not substantiated by the current data. The study is further compromised by a lack of criteria defining the diagnosis of pseudarthrosis & by heterogeneity of the pt population.

(continued)

TABLE 1: Bone growth stimulators as an adjunct for lumbar fusion: summary of evidence* *(continued)*

Authors & Year	Level of Evidence	Description	Comments
Kucharzyk, 1999	IV	This retrospective review was intended to determine the efficacy of DCS in high-risk pts undergoing fusion. 65 pts treated w/ DCS were compared w/ an equal no. of pts not treated w/ DCS. Variation existed w/ respect to presenting diagnosis & no. of levels fused. Fusion status was determined through radiographs & CT (evaluated by an independent radiologist). Clinical status was determined w/ a modified Smiley-Webster scale. Follow-up was conducted at regular intervals, w/ an average follow-up of 3.8 yrs. The overall fusion rate was 95.6% in the DCS group & 87% in the control group. The rate of clinical success was greater in pts receiving DCS (91% vs 79%). These differences were statistically significant. In a subgroup analysis of Workers' Compensation pts, fusion success was observed in 93% receiving DCS & in 81% of controls. Clinical success in this group was 57% in pts receiving DCS & 46% in the no-DCS group. The authors concluded that DCS significantly improved fusion & clinical success.	Limited baseline demographic data are provided. Although baseline demographics were similar in the DCS and no-DCS groups, the authors failed to adequately analyze these data to determine if significant heterogeneity existed btwn groups. The average follow-up was excellent, w/ no pts lost to follow-up in either group. The study benefits from an independent review of imaging, w/ fusion criteria being well defined. A nonvalidated clinical outcomes instrument was used. The study was downgraded to Level IV evidence due to the limitations described.
Rogozinski & Rogozinski, 1996	IV	The objective of this retrospective review was to determine the efficacy of DCS in pts undergoing instrumented fusion. 94 pts (53 receiving DCS) w/ varying diagnoses & surgical procedures were included. 26 pts were randomly assigned to treatment & control groups. Treating surgeons evaluated static & dynamic radiographs to determine fusion status. Follow-up was performed at 3, 6, & 12 mos after surgery. Fusion was demonstrated in 96% of pts receiving DCS while only 85% of control pts achieved solid fusion. The authors concluded that DCS can improve fusion rates in pts undergoing fusion, including high-risk pts (smokers, pts undergoing multilevel fusions).	Potential for selection bias exists due to the heterogeneous population of pts, w/ respect to both diagnosis & no. of levels fused. The follow-up period was limited to 12 mos. The authors failed to perform a blinded assessment of fusion, w/ questionable means of fusion assessment in presence of instrumentation. No clinical data were provided. Description of the statistical analysis was not provided. Due to these limitations, the study was downgraded to Level IV evidence in support of DCS.

* BMP = bone morphogenetic protein; DC = direct current; DCS = DC stimulation; DPQ = Dallas Pain Questionnaire; LBPRS = Low Back Pain Rating Scale; PEMFS = pulsed electromagnetic field stimulation; pt = patient; SF-36 = 36-Item Short Form Health Survey.

low—33% and 32%, respectively. The authors concluded that the utilization of DCS had no impact on fusion rate. There was a poor correlation between the observations made from CT images and plain radiographs, although a solid fusion, as defined by CT, resulted in better functional outcome and less pain[2] (Table 1). The final study from this series demonstrated, through the use of dual energy x-ray absorptiometry (DEXA), that the application of DCS had no impact on the bone mineral density of the fusion mass.[4] These investigations suffer from the same limitations as the first study in this series. The blinded radiographic assessment of the CT imaging strengthens the observations and conclusions regarding the impact of DCS on fusion rate. Although the authors claim that there was an 89% follow-up rate at 2 years, this was calculated from the 95 patients undergoing imaging at 1 year and not from the original 107 patients randomized at the onset of the study. Like the original study, this investigation was downgraded and provides Level II evidence against the utility of DCS to enhance the fusion rate for noninstrumented lumbar fusion. In a follow-up study, Andersen et al. investigated the impact of DCS on the quality of fusion formation by examining 80 of the original 107 patients with DEXA at 1 year after surgery. No significant difference in bone mineral density was observed between the 3 treatment groups.

Rogozinski et al. conducted a prospective, nonrandomized trial comparing radiographic outcome in 31 patients with the diagnosis of degenerative disc disease, who underwent 1- to 3-level instrumented PLF supplemented with either bone morphogenetic protein (BMP) or an implanted DC stimulator.[20] Fusion status was assessed using plain radiographs and/or CT, and pain status was determined through a 10-point visual analog scale (VAS). The BMP cohort demonstrated a 100% fusion rate, while the stimulator group demonstrated a 93.4% fusion rate. The BMP cohort was considered to achieve more robust fusion at a faster rate than the stimulator cohort. Pain improved in both cohorts. The authors concluded that the use of BMP led to more rapid graft maturation and a more robust fusion compared with fusions supplemented with an internal stimulator. The actual treatment effect of DCS compared with traditional fusion techniques cannot be determined from this study because all patients received some form of fusion supplement, but the fusion rate observed in the DCS cohort is comparable to previously reported rates of fusion for similar patients without DCS. This investigation also suffers from major limitations with respect to study design, including a small, heterogeneous patient cohort, lack of inclusion and exclusion criteria, and heterogeneous surgical treatments. Only 9 patients were available for CT imaging at 2 years after surgery (79% lost to follow-up), the assessment of the images was performed by the treating surgeon, and the criteria for fusion were not defined. This study was therefore downgraded to Level III evidence, although one may consider it simply a case series with respect to the DCS data (Table 1).

Two additional studies have also demonstrated a positive impact of DCS on fusion formation. Kucharzyk performed a retrospective review of 130 cases involving patients undergoing lumbar fusion with (n = 65) and without (n = 65) placement of DCS.[13] Fusion status was determined through both CT images and plain radiographs. The average follow-up was 3.8 years. The fusion rate in the DCS cohort was 95.6%, while the rate in the control group was 87%. Clinical success, utilizing a nonvalidated outcome measure, was superior in the DCS group. Rogozinski and Rogozinski also performed a retrospective review of 94 cases, with 53 of the patients receiving a DCS, and observed a fusion rate of 96% in the DCS cohort and 85% in the control arm.[19] Both of these studies suffer from a heterogeneous population of patients, limited baseline demographic data, and either failure to report clinical outcome or use of a nonvalidated instrument. Due to these limitations these studies are downgraded to Level IV evidence in support of the use of DCS with lumbar fusions.

Simmons et al. published a case series involving 100 patients with a mean age of 43.3 years who presented with pseudarthrosis after an attempt at single- or multilevel lumbar fusions and were treated with pulsed electromagnetic field stimulation (PEMFS).[21] Pseudarthrosis was confirmed by the presence of motion on dynamic imaging and the lack of visible bone healing on CT, MRI, or radiographic images. Twenty-five investigators from multiple institutions enrolled the 100 patients, who received PEMFS for at least 90 days. The investigators as well as a blinded radiologist performed radiographic evaluation of fusion. If there was disagreement among reviewers, an independent evaluation was performed by a blinded orthopedist. A solid fusion was defined as 50% or more assimilation of the graft based on radiographic imaging; the specific imaging technique was not defined. Clinical outcome was rated as excellent, good, fair, or poor, based on patients' reported pain intensity, medication usage, and return to work. The fusion success rate was 67%, and 63% of the patients with successful fusion demonstrated an excellent or good outcome. Only 30% of patients with persistent pseudarthrosis had an excellent or good outcome. The authors concluded that PEMFS was an effective alternative to revision surgery for patients presenting with pseudarthrosis. Although this study provides evidence that the utilization of PEMFS is a feasible intervention for the management of pseudarthrosis, the true treatment effect cannot be determined due to the study design and lack of an adequate control group. The authors also fail to define the criteria used to diagnose pseudarthrosis and included a heterogeneous population of patients. The study therefore provides at best Level IV evidence in support of PEMFS for treatment of pseudarthrosis (Table 1).

Summary

Based on the recommendations from the original guidelines, both DCS and CCES may be considered in patients at high risk for pseudarthrosis who are undergoing PLF, while PEMFS may be considered in a similar patient population undergoing an interbody fusion. Since the publication of the previous guidelines, there have been few clinical trials that provide further insight into the clinical utility of bone growth stimulation. The current data do not contribute to the previous recommendations.

The few studies that have investigated the use of bone growth stimulators have methodological flaws that compromise the conclusions and prohibit the formulation of strong recommendations. Based on a single Level II study, there is a suggestion that the use of DCS in patients over 60 years of age may provide a clinical benefit; however, this benefit was only observed in a subset of measures from a single outcome instrument and therefore is considered a weak correlation. This potential beneficial effect is further weakened by the fact that DCS did not have a positive impact on the fusion rate or quality in the same patient population. The weak correlation to clinical outcome may therefore be an artifact of the flawed study design or simply due to chance. Since the intended purpose of DCS was not supported by the authors' observations, the routine use of DCS in patients over 60 years of age undergoing a noninstrumented fusion was not recommended.

The second recommendation supports the use of PEMFS in patients suffering from a pseudarthrosis, but no comment can be made regarding the routine use of PEMFS. Due to the noninvasive nature of PEMFS, its application appears to be relatively benign with few drawbacks; however, in today's medical climate one cannot ignore the costs associated with an intervention that has not been proven to provide definitive benefit. Unfortunately, the quality of the current literature does not help to address these concerns.

Key Issues for Future Investigation

The impact of bone growth stimulators on fusion rates is likely to be minimal, and this makes it difficult to conduct a clinical trial to determine the actual treatment effect and/or compare the efficacy of different types of stimulators. Given the noninvasive nature of PEMFS, a well-designed randomized controlled trial is feasible. Such a study would, however, require an exceedingly large number of patients to demonstrate the difference in treatment effect. Nevertheless, such information would prove valuable, not only from a clinical perspective, but also for effective cost analysis, which ultimately may be the more relevant issue in today's medical climate. Utilization of a prospective patient registry may also provide relevant information by identifying specific patient populations that would benefit from any advantage provided by fusion enhancers, such as bone growth stimulators.

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., 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 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: Kaiser. 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: Kaiser. Study supervision: Kaiser.

References

1. Agarwal R, Williams K, Umscheid CA, Welch WC: Osteoinductive bone graft substitutes for lumbar fusion: a systematic review. Clinical article. **J Neurosurg Spine 11:**729–740, 2009
2. Andersen T, Christensen FB, Egund N, Ernst C, Fruensgaard S, Østergaard J, et al: The effect of electrical stimulation on lumbar spinal fusion in older patients: a randomized, controlled, multi-center trial. Part 2: fusion rates. **Spine (Phila Pa 1976) 34:**2248–2253, 2009
3. Andersen T, Christensen FB, Ernst C, Fruensgaard S, Østergaard J, Andersen JL, et al: The effect of electrical stimulation on lumbar spinal fusion in older patients: a randomized, controlled, multi-center trial. Part 1: functional outcome. **Spine (Phila Pa 1976) 34:**2241–2247, 2009
4. Andersen T, Christensen FB, Langdahl BL, Ernst C, Fruensgaard S, Ostergaard J, et al: Fusion mass bone quality after uninstrumented spinal fusion in older patients. **Eur Spine J 19:**2200–2208, 2010
5. Andersen T, Christensen FB, Niedermann B, Helmig P, Høy K, Hansen ES, et al: Impact of instrumentation in lumbar spinal fusion in elderly patients: 71 patients followed for 2-7 years. **Acta Orthop 80:**445–450, 2009
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. Dawson E, Bae HW, Burkus JK, Stambough JL, Glassman SD: Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. **J Bone Joint Surg Am 91:**1604–1613, 2009
8. Dimar JR II, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY: Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. **J Bone Joint Surg Am 91:**1377–1386, 2009
9. Dimar JR II, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY: Two-year fusion and clinical outcomes in 224 patients treated with a single-level instrumented postero-

lateral fusion with iliac crest bone graft. **Spine J 9:**880–885, 2009
10. Dwyer AF: The use of electrical current stimulation in spinal fusion. **Orthop Clin North Am 6:**265–273, 1975
11. Fritzell P, Hägg O, Wessberg P, Nordwall A: Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish Lumbar Spine Study Group. **Spine (Phila Pa 1976) 27:**1131–1141, 2002
12. Ghogawala Z, Benzel EC, Amin-Hanjani S, Barker FG II, Harrington JF, Magge SN, et al: Prospective outcomes evaluation after decompression with or without instrumented fusion for lumbar stenosis and degenerative Grade I spondylolisthesis. **J Neurosurg Spine 1:**267–272, 2004
13. Kucharzyk DW: A controlled prospective outcome study of implantable electrical stimulation with spinal instrumentation in a high-risk spinal fusion population. **Spine (Phila Pa 1976) 24:**465–469, 1999
14. Lawlis GF, Cuencas R, Selby D, McCoy CE: The development of the Dallas Pain Questionnaire. An assessment of the impact of spinal pain on behavior. **Spine (Phila Pa 1976) 14:**511–516, 1989
15. Oishi M, Onesti ST: Electrical bone graft stimulation for spinal fusion: a review. **Neurosurgery 47:**1041–1056, 2000
16. 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 1: introduction and methodology. **J Neurosurg Spine 2:**637–638, 2005
17. 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 12: pedicle screw fixation as an adjunct to posterolateral fusion for low-back pain. **J Neurosurg Spine 2:**700–706, 2005
18. 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 17: bone growth stimulators and lumbar fusion. **J Neurosurg Spine 2:**737–740, 2005
19. Rogozinski A, Rogozinski C: Efficacy of implanted bone growth stimulation in instrumented lumbosacral spinal fusion. **Spine (Phila Pa 1976) 21:**2479–2483, 1996
20. Rogozinski A, Rogozinski C, Cloud G: Accelerating autograft maturation in instrumented posterolateral lumbar spinal fusions without donor site morbidity. **Orthopedics 32:**809, 2009
21. Simmons JW Jr, Mooney V, Thacker I: Pseudarthrosis after lumbar spine fusion: nonoperative salvage with pulsed electromagnetic fields. **Am J Orthop 33:**27–30, 2004
22. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, et al: Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. Four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. **J Bone Joint Surg Am 91:**1295–1304, 2009

Manuscript submitted March 27, 2014.
Accepted April 9, 2014.
Please include this information when citing this paper: DOI: 10.3171/2014.4.SPINE14326.
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.