



**CONGRESS OF NEUROLOGICAL SURGEONS SYSTEMATIC REVIEW AND
EVIDENCE-BASED GUIDELINES ON THE EVALUATION AND TREATMENT
OF PATIENTS WITH THORACOLUMBAR SPINE TRAUMA:
RADIOLOGICAL EVALUATION**

Sponsored by: Congress of Neurological Surgeons and the Section on Disorders of the Spine and Peripheral Nerves in collaboration with the Section on Neurotrauma and Critical Care

Endorsed by: Joint Guidelines Committee of the American Association of Neurological Surgeons (AANS) and the Congress of Neurological Surgeons (CNS)

Sheeraz Qureshi, MD, MBA,¹ Sanjay S. Dhall, MD,² Paul A. Anderson, MD,³ Paul M. Arnold, MD,⁴ John H. Chi, MD, MPH,⁵ Andrew T. Dailey, MD,⁶ Kurt M. Eichholz, MD,⁷ James S. Harrop, MD,⁸ Daniel J. Hoh, MD,⁹ Craig H. Rabb, MD,¹⁰ P. B. Raksin, MD,¹¹ Michael G. Kaiser, MD,¹² and John E. O'Toole, MD, MS¹³

1. Department of Orthopaedic Surgery, Weill Cornell Medical College, New York, New York

2. Department of Neurological Surgery, University of California, San Francisco, San Francisco, California
3. Department of Orthopedics and Rehabilitation, University of Wisconsin, Madison, Wisconsin
4. Department of Neurosurgery, University of Kansas School of Medicine, Kansas City, Kansas
5. Department of Neurosurgery, Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts
6. Department of Neurosurgery, University of Utah, Salt Lake City, Utah
7. St. Louis Minimally Invasive Spine Center, St. Louis, Missouri
8. Departments of Neurological Surgery and Orthopedic Surgery, Thomas Jefferson University, Philadelphia, Pennsylvania
9. Lillian S. Wells Department of Neurological Surgery, University of Florida, Gainesville, Florida
10. Department of Neurosurgery, University of Utah, Salt Lake City, Utah
11. Division of Neurosurgery, John H. Stroger, Jr. Hospital of Cook County and Department of Neurological Surgery, Rush University Medical Center, Chicago, Illinois
12. Department of Neurosurgery, Columbia University, New York, New York
13. Department of Neurological Surgery, Rush University Medical Center, Chicago, Illinois

Correspondence:

Sheeraz Qureshi, MD, MBA

Hospital for Special Surgery

535 East 70th Street

New York, NY 10021

Email: sheerazqureshimd@gmail.com

Keywords: Clinical guidelines, diagnostic imaging, magnetic resonance imaging (MRI), spinal injury, spine fracture, thoracolumbar, risk assessment

Abbreviations

AANS – American Association of Neurological Surgeons

AIS – American Spinal Injury Association Impairment Scale

AP – antero-posterior

AO – Arbeitsgemeinschaft für Osteosynthesefragen (Association for the Study of Internal Fixation)

BASIC – Brain and Spinal Injury Center

CT – computed tomography

CNS – Congress of Neurological Surgeons

JGC – Joint Guidelines Committee

MRI – magnetic resonance imaging

PLC – posterior ligamentous complex

TL – thoracolumbar

TLICS – thoracolumbar injury classification and severity

No part of this article has been published or submitted for publication elsewhere.

ABSTRACT

Background: Radiologic evaluation of traumatic thoracolumbar fractures is used to classify the injury and determine the optimal treatment plan. Currently, there remains a lack of consensus regarding appropriate radiologic protocol. Most clinicians use a combination of plain radiographs, three-dimensional computed tomography (CT) with reconstructions, and magnetic resonance imaging (MRI).

Objective: The purpose of this evidence-based guidelines review was to determine: (1) whether the use of MRI to identify ligamentous integrity predicted the need for surgical intervention; and (2) if there are any radiologic findings that can assist in predicting clinical outcomes.

Methods: A systematic review of the literature was performed using the National Library of Medicine/PubMed database and the Cochrane Library for studies relevant to thoracolumbar trauma. Clinical studies specifically addressing the radiologic evaluation of thoracolumbar spine trauma were selected for review.

Results: Two of 2278 studies met inclusion criteria for review. One retrospective review (level III) and one prospective cohort (level III) provided evidence that the addition of an MRI scan in acute thoracic and thoracolumbar trauma can predict the need for surgical intervention. There was insufficient evidence that MRI can help predict clinical outcomes in patients with acute traumatic thoracic and thoracolumbar spine injuries.

Conclusion: This evidence-based guideline provides a grade B recommendation that radiologic findings in patients with acute thoracic or thoracolumbar spine trauma can

predict the need for surgical intervention. This evidence-based guideline provides a grade insufficient recommendation that there is insufficient evidence to determine if radiographic findings can assist in predicting clinical outcomes in patients with acute thoracic and thoracolumbar spine injuries.

RECOMMENDATIONS

Question 1

Are there radiographic findings in patients with traumatic thoracolumbar fractures that can predict the need for surgical intervention?

Recommendation 1

Because magnetic resonance imaging has been shown to influence the management of up to 25% of patients with thoracolumbar fractures, providers may use magnetic resonance imaging to assess posterior ligamentous complex integrity, when determining the need for surgery.

Strength of Recommendation: Grade B

Question 2

Are there radiographic findings in patients with traumatic thoracolumbar fractures that can assist in predicting clinical outcomes?

Recommendation 2

Due to a paucity of published studies, there is insufficient evidence that radiographic findings can be used as predictors of clinical outcomes in thoracolumbar fractures.

Strength of Recommendation: Grade Insufficient

INTRODUCTION

Goals and Rationale

This clinical guideline was created to improve patient care by outlining the appropriate information-gathering and decision-making processes involved in the evaluation and treatment of patients with thoracolumbar spine trauma. Spinal surgical care is provided in many different settings by many different providers. This guideline was created as an educational tool to guide qualified physicians through a series of diagnostic and treatment decisions to improve the quality and efficiency of care.

Fractures of the thoracic and lumbar region constitute a spectrum of injuries ranging from simple nondisplaced fractures to complex fracture dislocations.¹ Anatomically and functionally, the thoracic spine is rigid with coronal-oriented facet joints and thin intervertebral discs, while the lumbar spine is relatively flexible, with sagittal-oriented facet joints and thicker discs. The thoracolumbar junction, being uniquely positioned between the rigid thoracic spine and the flexible lumbar spine, is subject to significant biomechanical stress. Fractures of this region are the most common injuries of the vertebral column.²

Fractures of the thoracolumbar spine are often unstable, resulting in significant disability, deformity, and neurologic deficit. Standard classification systems have been devised to help with communication and guide treatment. These classification systems are based on injury mechanism, fracture morphology, injury to the posterior ligamentous complex (PLC), and neurologic deficit. While plain radiographs are often obtained, a computed

tomography (CT) scan is generally required to provide information on the extent of bony injury and a magnetic resonance imaging (MRI) scan is required to assess the spinal cord and soft tissue structures.

Standard radiographic evaluation most often includes anteroposterior (AP) and lateral x-rays. These images are used to evaluate spinal alignment, rotatory or translational instability, loss of vertebral body height, and widening of interpedicular or interspinous distance.³⁻⁵ A CT scan is used to characterize the fracture and assess the degree of spinal canal compromise. Studies have found that up to 25% of burst fractures are incorrectly diagnosed as compression fractures when plain x-rays alone are used.⁶ MRI is used to obtain information on spinal cord or nerve root injury and the presence of spinal cord edema or epidural hematoma.⁷ MRI also provides information on injuries to the discs and PLC, and a sagittal screening MRI identifies the presence of noncontiguous spine injuries that can be seen in nearly 25% of cases.^{8,9}

Despite extensive studies on thoracic and thoracolumbar fractures, several areas of controversy still exist. One of these areas is the impact of radiological findings on treatment decision and patient outcome. The purpose of this evidence-based guideline is to provide information on whether there are radiological findings that can predict the need for surgical intervention or assist in predicting patient outcomes.

METHODS

The guidelines task force initiated a systematic review of the literature relevant to the diagnosis and treatment of patients with thoracolumbar trauma. Through objective evaluation of the evidence and transparency in the process of making recommendations, this evidence-based clinical practice guideline was developed for the diagnosis and treatment of adult patients with thoracolumbar injuries. These guidelines are developed for educational purposes to assist practitioners in their clinical decision-making processes. Additional information about the methods used in this systematic review is provided in the [introduction and methodology chapter](#).

Literature Search

The task force members identified search terms/parameters, and a medical librarian implemented the literature search, consistent with the literature search protocol (see Appendix I), searching the National Library of Medicine PubMed database and the Cochrane Library (which included the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effect, the Cochrane Central Register of Controlled Trials, the Health Technology Assessment Database, and the National Health Service Economic Evaluation Database) for the period from January 1, 1946, to March 31, 2015, using the search strategies provided in Appendix I.

RESULTS

The literature search yielded 2278 abstracts. Task force members reviewed all abstracts yielded from the literature search and identified the literature for full text review and extraction, addressing the clinical questions, in accordance with the Literature Search Protocol (Appendix I). Task force members identified the best research evidence available to answer the targeted clinical questions. When level I, II, or III literature was available to answer specific questions, the task force did not review level IV studies.

The task force selected 126 articles for full-text review. Of these, 124 were rejected for not meeting inclusion criteria or for being off topic. Three were selected for inclusion in this systematic review (Appendix II).

Inclusion/Exclusion Criteria

Articles were retrieved and included only if they met specific inclusion/exclusion criteria. These criteria were also applied to articles provided by guideline task force members who supplemented the electronic database searches with articles from their own files. To reduce bias, these criteria were specified before conducting the literature searches.

Articles that did not meet the following criteria, for the purposes of this evidence-based clinical practice guideline, were excluded. To be included as evidence in the guideline, an article had to be a report of a study that:

- Investigated patients with thoracolumbar injuries;
- Included patients ≥ 18 years of age;
- Enrolled $\geq 80\%$ of thoracolumbar injuries (studies with mixed patient populations

were included if they reported results separately for each group/patient population);

- Was a full article report of a clinical study;
- Was not an internal medical records review, meeting abstract, historical article, editorial, letter, or commentary;
- Appeared in a peer-reviewed publication or a registry report;
- Enrolled ≥ 10 patients per arm per intervention (20 total) for each outcome;
- Included only human subjects;
- Was published in or after 1946;
- Quantitatively presented results;
- Was not an in vitro study;
- Was not a biomechanical study;
- Was not performed on cadavers;
- Was published in English;
- Was not a systematic review, meta-analysis, or guideline developed by others*.
- Was a case series (therapeutic study) where higher level evidence exists.

Rating Quality of Evidence

The guideline task force used a modified version of the North American Spine Society's (NASS) evidence-based guideline development methodology. The NASS methodology

*The guideline task force did not include systematic reviews, guidelines, or meta-analyses conducted by others. These documents are developed using different inclusion criteria than those specified in this guideline; therefore, they may include studies that do not meet the inclusion criteria specific in this guideline. In cases where these types of documents' abstract suggested relevance to the guideline's recommendations, the task force searched their bibliographies for additional studies.

uses standardized levels of evidence (Appendix III) and grades of recommendation (Appendix IV) to assist practitioners in easily understanding the strength of the evidence and recommendations within the guidelines. The levels of evidence range from level I (high-quality randomized controlled trial) to level IV (case series). Grades of recommendation indicate the strength of the recommendations made in the guideline based on the quality of the literature. Levels of evidence have specific criteria and are assigned to studies before developing recommendations. Recommendations are then graded based upon the level of evidence. To better understand how levels of evidence inform the grades of recommendation and the standard nomenclature used within the recommendations, see Appendix IV.

Guideline recommendations were written using a standard language that indicates the strength of the recommendation. “A” recommendations indicate a test or intervention is “recommended”; “B” recommendations “suggest” a test or intervention; “C” recommendations indicate a test or intervention or “is an option.” “Insufficient evidence” statements clearly indicate that “there is insufficient evidence to make a recommendation for or against” a test or intervention. Task force consensus statements clearly state that “in the absence of reliable evidence, it is the task force’s opinion that” a test or intervention may be considered. Both the levels of evidence assigned to each study and the grades of each recommendation were arrived at by consensus of the workgroup employing up to three rounds of voting when necessary.

In evaluating studies as to levels of evidence for this guideline, the study design was interpreted as establishing only a potential level of evidence. As an example, a therapeutic study designed as a randomized controlled trial would be considered a potential level I study. The study would then be further analyzed as to how well the study design was implemented and significant shortcomings in the execution of the study would be used to downgrade the levels of evidence for the study's conclusions (see Appendix V for additional information and criteria).

Revision Plans

In accordance with the Institute of Medicine's standards for developing clinical practice guidelines and criteria specified by the National Guideline Clearinghouse, the task force will monitor related publications following the release of this document and will revise the entire document and/or specific sections "if new evidence shows that a recommended intervention causes previously unknown substantial harm; that a new intervention is significantly superior to a previously recommended intervention from an efficacy or harms perspective; or that a recommendation can be applied to new populations."¹⁰ In addition, the task force will confirm within five years from the date of publication that the content reflects current clinical practice and the available technologies for the evaluation and treatment for patients with thoracolumbar trauma.

DISCUSSION

Question 1

Are there radiographic findings in patients with traumatic thoracolumbar fractures that can predict the need for surgical intervention?

Recommendation 1

Because magnetic resonance imaging has been shown to influence the management of up to 25% of patients with thoracolumbar fractures, providers may use magnetic resonance imaging to assess posterior ligamentous complex integrity, when determining the need for surgery.

Strength of Recommendation: Grade B

Winklhofer et al¹¹ evaluated the influence of additional MRI compared to CT alone for the classification of traumatic spinal injuries using the AO system and the Thoracolumbar Injury Classification and Severity (TLICS) scale. The authors retrospectively reviewed the images of 100 consecutive patients with at least one fracture on CT with regard to AO and TLICS classification systems in 2 steps. The first step was to analyze the initial CT scan, and the second step was to analyze the CT scan and MRI together 6 weeks later. Statistical analysis was performed to identify changes in the number of fractures and ligamentous lesions detected and their corresponding classification.

In this study, 162 fractures were identified on the initial CT scan. Review of CT scan and MRI together revealed a total of 196 fractures. The AO classification changed in 31%, and the TLICS classification changed in 33% of patients after MRI review. Based on the evaluation of the CT and MRI together, the TLICS value changed from values <5

(indication for conservative therapy) to values ≥ 5 (indication for surgical therapy) in 24% of patients.

Because of its heterogeneous patient population, this retrospective study was downgraded from level II evidence to level III evidence that the addition of an MRI in patients with thoracolumbar fractures can provide findings that can help predict the need for surgical intervention.

Pizones et al¹² conducted a prospective study to analyze the usefulness of MRI in fracture diagnosis and its influence on treatment decision making. Acute traumatic thoracolumbar fractures in 33 patients were initially classified using X-ray and CT scan following the AO classification. A selective MRI was then performed and the fractures were classified according to the TLICS system and reclassified following the AO system. The authors analyzed diagnostic changes, occult fractures, and differences in treatment decision making before and after the MRI.

Forty-one fractures were initially diagnosed using plain x-rays and CT scans. Following the MRI, 50 fractures and 9 vertebral contusions were diagnosed. The authors reported that the addition of an MRI modified the diagnosis in 40% of patients, the classification of fracture pattern in 24% of fractures, and the therapeutic management in 16% of patients.

This study by Pizones et al¹² was graded level III evidence due to its small sample size and lack of consecutive patients. This study, like the Winklhofer et al study,¹¹ provides

level III evidence that the findings on an MRI can help predict the need for surgical intervention.

Question 2

Are there radiographic findings in patients with traumatic thoracolumbar fractures that can assist in predicting clinical outcomes?

Recommendation 2

Due to a paucity of published studies, there is insufficient evidence that radiographic findings can be used as predictors of clinical outcomes in thoracolumbar fractures.

Strength of Recommendation: Grade Insufficient

Future Research

Several gaps exist in the literature regarding the ability of radiologic studies to predict the need for surgery and clinical outcomes in patients with acute traumatic thoracolumbar spine injuries. Currently, there is only level III evidence based on a limited number of studies that MRI can help to predict the need for surgery and the severity of spinal cord injury. There is a need for studies that provide a higher level of evidence for each of these questions.

Furthermore, it remains difficult for clinicians using MRI to visualize with certainty a complete versus incomplete rupture of the PLC. In addition, current imaging technology is unable to assist in identifying those PLC ruptures that are self-healing versus those that

will require surgery to prevent collapse. Future radiologic studies that focus on the characteristics of PLC injuries will be valuable to fill the knowledge gap.

Currently, there is insufficient evidence that radiologic findings can be used to predict clinical outcomes. Additional research is needed to test the capacity of MRI scans and other imaging modalities to predict long-term outcome.

Conclusions

Two studies provide level III evidence that MRI scans in patients with acute thoracic and thoracolumbar spine trauma can impact classification of injury and the decision to proceed with surgical intervention. Winklhofer et al¹¹ found that the addition of an MRI resulted in improved fracture identification and change in AO classification and well as increase in TLICS score. Similarly, Pizones et al¹² reported that an MRI scan resulted in a change in diagnosis, classification, and therapeutic management.

With respect to radiologic findings assisting in the prediction of clinical outcomes, there is insufficient evidence that MRI is useful.

The existing evidence suggests that MRI of patients with thoracolumbar spinal trauma improves the detection of fractures and soft tissue compared with CT alone and changes the overall trauma classification. MRI is a useful tool in the evaluation of acute thoracolumbar fractures as it allows for better visualization of the posterior ligamentous

complex integrity and of the levels involved, offering additional information compared to traditional diagnostic tools.

Potential Conflicts of Interest

The task force members were required to report all possible conflicts of interest (COIs) prior to beginning work on the guideline, using the COI disclosure form of the AANS/CNS Joint Guidelines Committee, including potential COIs that are unrelated to the topic of the guideline. The CNS Guidelines Committee and Guideline Task Force Chairs reviewed the disclosures and either approved or disapproved the nomination. The CNS Guidelines Committee and Guideline Task Force Chairs are given latitude to approve nominations of Task Force members with possible conflicts and address this by restricting the writing and reviewing privileges of that person to topics unrelated to the possible COIs. The conflict of interest findings are provided in detail in the companion [introduction and methods manuscript](#).

Disclaimer of Liability

This clinical systematic review and evidence-based guideline was developed by a multidisciplinary physician volunteer task force and serves as an educational tool designed to provide an accurate review of the subject matter covered. These guidelines are disseminated with the understanding that the recommendations by the authors and consultants who have collaborated in their development are not meant to replace the individualized care and treatment advice from a patient's physician(s). If medical advice or assistance is required, the services of a competent physician should be sought. The proposals contained in these guidelines may not be suitable for use in all circumstances.

The choice to implement any particular recommendation contained in these guidelines must be made by a managing physician in light of the situation in each particular patient and on the basis of existing resources.

Disclosures

These evidence-based clinical practice guidelines were funded exclusively by the Congress of Neurological Surgeons and the Section on Disorders of the Spine and Peripheral Nerves in collaboration with the Section on Neurotrauma and Critical Care, which received no funding from outside commercial sources to support the development of this document.

Acknowledgments

The guidelines task force would like to acknowledge the CNS Guidelines Committee for their contributions throughout the development of the guideline and the AANS/CNS Joint Guidelines Committee for their review, comments, and suggestions throughout peer review, as well as the contributions of Trish Rehring, MPH, CHES, Senior Manager of Clinical Practice Guidelines for the CNS, and Mary Bodach, MLIS, Guidelines Specialist and Medical Librarian for assistance with the literature searches. Throughout the review process the reviewers and authors were blinded from one another. At this time, the guidelines task force would like to acknowledge the following individual peer reviewers for their contributions: Maya Babu, MD, MBA, Greg Hawryluk, MD, PhD, Steven Kalkanis, MD, Yi Lu, MD, PhD, Jeffrey J. Olson, MD, Martina Stippler, MD, Cheerag Upadhyaya, MD, MSc, and Robert Whitmore, MD.

REFERENCES

1. Wood KB, Li W, Lebl DR, Ploumis A. Management of thoracolumbar spine fractures. *Spine J* 2014;14:145-164.
2. el-Khoury GY, Whitten CG. Trauma to the upper thoracic spine: Anatomy, biomechanics, and unique imaging features. *AJR Am J Roentgenol* 1993;160:95-102.
3. Keene JS. Radiographic evaluation of thoracolumbar fractures. *Clin Orthop Relat Res* 1984;189:58-64.
4. Dalinka MK, Kessler H, Weiss M. The radiographic evaluation of spinal trauma. *Emerg Med Clin North Am* 1985;3:475-490.
5. Harris Jr JH. Radiographic evaluation of spinal trauma. *Orthop Clin North Am* 1986;17:75-86.
6. Ballock RT, Mackersie R, Abitbol JJ, Cervilla V, Resnick D, Garfin SR. Can burst fractures be predicted from plain radiographs? *J Bone Joint Surg Br* 1992;74:147-150.
7. Tarr RW, Drolshagen LF, Kerner TC, Allen JH, Partain CL, James Jr AE. MR imaging of recent spinal trauma. *J Comput Assist Tomogr* 1987;11:412-417.
8. Korres DS, Boscainos PJ, Papagelopoulos PJ, Psycharis I, Goudelis G, Nikolopoulos K. Multiple level noncontiguous fractures of the spine. *Clin Orthop Relat Res* 2003;411:95-102.
9. Henderson RL, Reid DC, Saboe LA. Multiple noncontiguous spine fractures. *Spine (Phila Pa 1976)* 1991;16:128-131.
10. Ransohoff DF, Pignone M, Sox HC. How to decide whether a clinical practice guideline is trustworthy. *JAMA* 2013;309:139-140.

11. Winklhofer S, Thekkumthala-Sommer M, Schmidt D, et al. Magnetic resonance imaging frequently changes classification of acute traumatic thoracolumbar spine injuries. *Skeletal Radiol* 2013;42:779-786.
12. Pizones J, Izquierdo E, Alvarez P, et al. Impact of magnetic resonance imaging on decision making for thoracolumbar traumatic fracture diagnosis and treatment. *Eur Spine J* 2011;20(suppl 3):390-396.

Appendix I. Literature Searches

Search Strategies

PubMed

1. Lumbar vertebrae [MeSH] OR Thoracic vertebrae [MeSH]
2. Spinal Injuries [MeSH] OR Spinal Cord Injuries [MeSH]
3. #1 AND #2
4. Thoracolumbar [TIAB] OR thoraco-lumbar [TIAB] OR thoraco lumbar [TIAB]
OR burst [Title]
5. Injur* [TIAB] OR trauma* [TIAB] OR fractur* [TIAB] OR dislocation* [TIAB]
6. #4 AND #5
7. Lumbar vertebrae/injuries [MeSH] OR Thoracic vertebrae/injuries [MeSH]
8. #3 OR #6 OR #7
9. Diagnostic Imaging [MeSH] OR radiography [SH]
10. “MRI” OR “magnetic resonance imaging” OR “CT scan” OR “CT” OR
“computed tomography” OR x-ray* OR Imag* OR radiograph* [TIAB]
11. #9 OR #10
12. #8 AND #11
13. (animal [MeSH] NOT human [MeSH]) OR cadaver [MeSH] OR cadaver* [Titl]
OR comment [PT] OR letter [PT] OR editorial [PT] OR addresses [PT] OR news
[PT] OR “newspaper article” [PT] OR case reports [PT]
14. #12 NOT #13
15. osteoporosis [MH] OR osteoporotic fractures [MH] OR osteoporo* [TITLE] OR
spinal neoplasms [MH] OR tumor* [TITLE] OR tumour* [TITLE] OR malignan*

[TITLE]

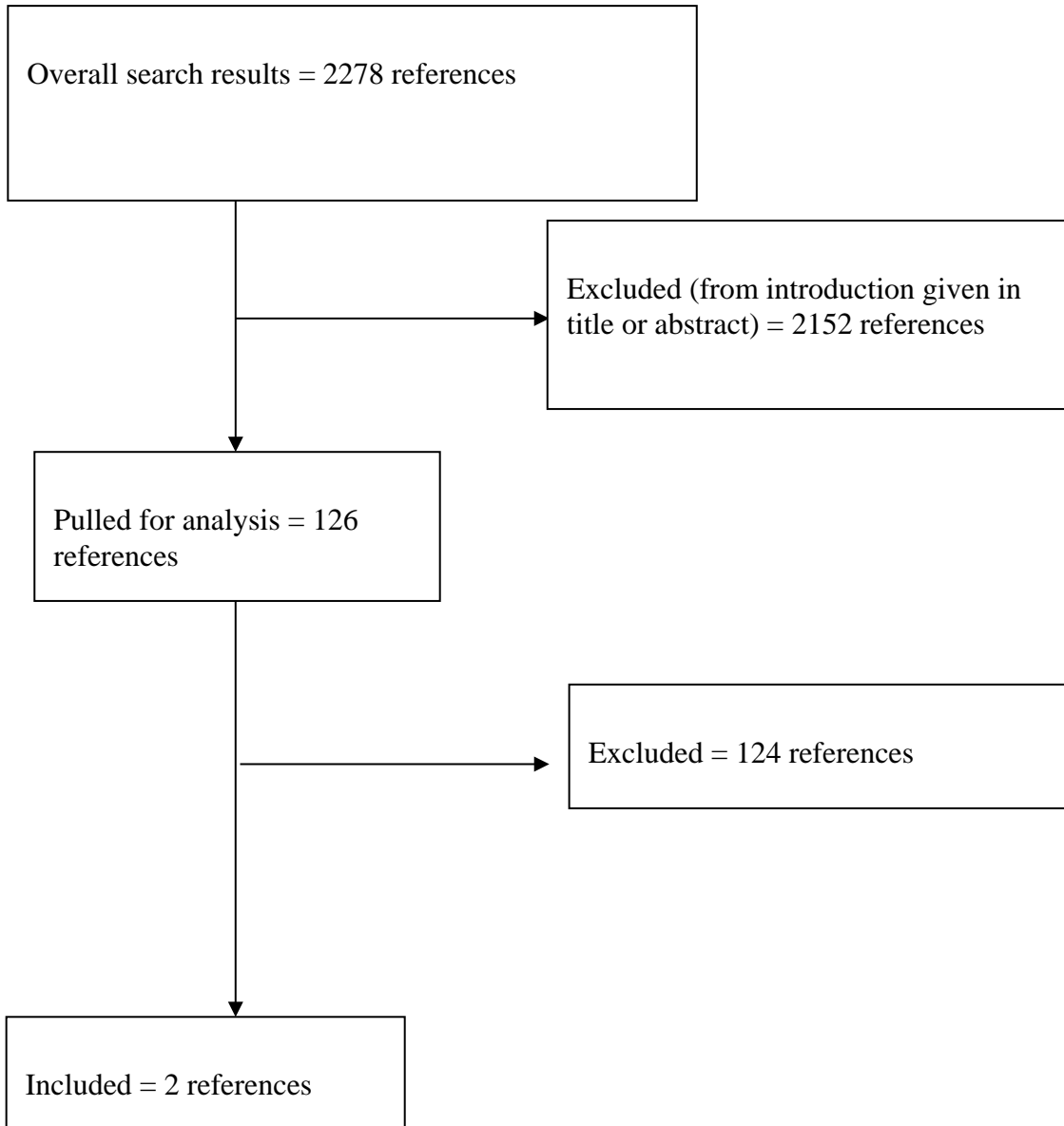
16. #14 NOT #15

17. #16 AND English [Lang]

Cochrane Library

1. Lumbar vertebrae: MeSH descriptor, explode all trees
2. Thoracic vertebrae: MeSH descriptor, explode all trees
3. #1 OR #2
4. Spinal Injuries: MeSH descriptor
5. Spinal Cord Injuries: MeSH descriptor
6. #4 OR #5
7. #3 AND #6
8. (Thoracolumbar OR thoraco-lumbar OR thoraco lumbar OR burst) NEAR/4
(Injur* OR trauma* OR fractur* OR dislocation*):ti,ab,kw
9. Lumbar vertebrae/injuries: MeSH descriptor, explode all trees
10. Thoracic vertebrae/injuries: MeSH descriptor, explode all trees
11. #9 OR #10
12. #7 OR #8 OR #11
13. mh osteoporosis or mh osteoporotic fractures or mh spinal neoplasms
14. osteopor* or tumor* or malignan*:ti
15. #13 OR #14
16. #12 NOT #15

Appendix II. Article Inclusions and Exclusions



Appendix III. Rating Evidence Quality

Levels of Evidence for Primary Research Question^a

Types of studies				
	Therapeutic studies – Investigating the results of treatment	Prognostic studies – Investigating the effect of a patient characteristic on the outcome of disease	Diagnostic studies – Investigating a diagnostic test	Economic and decision analyses – Developing an economic or decision model
Level I	<ul style="list-style-type: none"> • High-quality randomized trial with statistically significant difference or no statistically significant difference but narrow confidence intervals • Systematic review^b of level I RCTs (and study results were homogenous^c) 	<ul style="list-style-type: none"> • High-quality prospective study^d (all patients were enrolled at the same point in their disease with $\geq 80\%$ follow-up of enrolled patients) • Systematic review^b of level I studies 	<ul style="list-style-type: none"> • Testing of previously developed diagnostic criteria on consecutive patients (with universally applied reference “gold” standard) • Systematic review^b of level I studies 	<ul style="list-style-type: none"> • Sensible costs and alternatives; values obtained from many studies; with multiway sensitivity analyses • Systematic review^b of level I studies

Level II	<ul style="list-style-type: none"> • Lesser quality RCT (e.g., $\leq 80\%$ follow-up, no blinding, or improper randomization) • Prospective^d comparative study^e • Systematic review^b of level II studies or level I studies with inconsistent results 	<ul style="list-style-type: none"> • Retrospective^f study • Untreated controls from an RCT • Lesser quality prospective study (e.g., patients enrolled at different points in their disease or $\leq 80\%$ follow-up) • Systematic review^b of level II studies 	<ul style="list-style-type: none"> • Development of diagnostic criteria on consecutive patients (with universally applied reference “gold” standard) • Systematic review^b of level II studies 	<ul style="list-style-type: none"> • Sensible costs and alternatives; values obtained from limited studies; with multiway sensitivity analyses • Systematic review^b of level II studies
Level III	<ul style="list-style-type: none"> • Case control study^g • Retrospective^f comparative study^e • Systematic review^b of level III studies 	<ul style="list-style-type: none"> • Case control study^g 	<ul style="list-style-type: none"> • Study of non consecutive patients; without consistently applied reference “gold” standard • Systematic review^b of level III studies 	<ul style="list-style-type: none"> • Analyses based on limited alternatives and costs; and poor estimates • Systematic review^b of level III studies
Level IV	Case series ^h	Case series	<ul style="list-style-type: none"> • Case-control study • Poor reference standard 	<ul style="list-style-type: none"> • Analyses with no sensitivity analyses

RCT, Randomized controlled trial.

^aA complete assessment of quality of individual studies requires critical appraisal of all aspects of the study design.

^bA combination of results from ≥ 2 previous studies.

^cStudies provided consistent results.

^dStudy was started before the first patient enrolled.

^ePatients treated one way (e.g., instrumented arthrodesis) compared with a group of patients treated in another way (e.g., uninstrumented arthrodesis) at the same institution.

^fThe study was started after the first patient enrolled.

^gPatients identified for the study based on their outcome, called “cases” (e.g., pseudoarthrosis) are compared to those who did not have outcome, called “controls” (e.g., successful fusion).

^hPatients treated one way with no comparison group of patients treated in another way.

Appendix IV. Linking Levels of Evidence to Grades of Recommendation

Grade of recommendation	Standard language	Levels of evidence	
A	Recommended	Two or more consistent level I studies	
B	Suggested	One level I study with additional supporting level II or III studies	Two or more consistent level II or III studies
C	Is an option	One level I, II, or III study with supporting level IV studies	Two or more consistent level IV studies
Insufficient (insufficient or conflicting evidence)	Insufficient evidence to make recommendation for or against	A single level I, II, III, or IV study without other supporting evidence	>1 study with inconsistent findings ^a

^aNote that in the presence of multiple consistent studies, and a single outlying, inconsistent study, the Grade of Recommendation will be based on the level of the consistent studies.

Appendix V. Criteria Grading the Evidence

The task force used the criteria provided below to identify the strengths and weaknesses of the studies included in this guideline. Studies containing deficiencies were downgraded one level (no further downgrading allowed, unless so severe that study had to be excluded). Studies with no deficiencies based on study design and contained clinical information that dramatically altered current medical perceptions of topic were upgraded.

1. Baseline study design (i.e. therapeutic, diagnostic, prognostic) determined to assign initial level of evidence.
2. Therapeutic studies reviewed for following deficiencies:
 - Failure to provide a power calculation for an RCT;
 - High degree of variance or heterogeneity in patient populations with respect to presenting diagnosis/demographics or treatments applied;
 - Less than 80% of patient follow-up;
 - Failure to utilize validated outcomes instrument;
 - No statistical analysis of results;
 - Cross over rate between treatment groups of greater than 20%;
 - Inadequate reporting of baseline demographic data;
 - Small patient cohorts (relative to observed effects);
 - Failure to describe method of randomization;
 - Failure to provide flowchart following patients through course of study (RCT);
 - Failure to account for patients lost to follow-up;
 - Lack of independent post-treatment assessment (e.g., clinical, fusion status, etc.);
 - Utilization of inferior control group:
 - Historical controls;

- Simultaneous application of intervention and control within same patient.
 - Failure to standardize surgical/intervention technique;
 - Inadequate radiographic technique to determine fusion status (e.g. – static radiographs for instrumented fusion).
 - If an RCT fails criteria specific to RCT (such as method randomization reported or improper, no power, greater than 20% crossover, if there is or is not post treatment assessment, inappropriate statistics, no baseline data, small cohorts, etc.), then it will be initially assigned to level II. Only if it further fails additional evaluation, can it be downgraded further to a level III.
3. Methodology of diagnostic studies reviewed for following deficiencies:
- Failure to determine specificity and sensitivity;
 - Failure to determine inter- and intra-observer reliability;
 - Failure to provide correlation coefficient in the form of kappa values.
4. Methodology of prognostic studies reviewed for following deficiencies:
- High degree of variance or heterogeneity in patient populations with respect to presenting diagnosis/demographics or treatments applied;
 - Failure to appropriately define and assess independent and dependent variables (e.g., failure to use validated outcome measures when available).

Appendix VI. Evidence Tables

Author, Year	Level of Evidence	Task Force Conclusions Relative to Question and Rationale for Evidence Grading
Pizones et al, ¹² 2011	III	This paper provides evidence that multidimensional analysis of magnetic resonance imaging predicts early impairment in thoracic and thoracolumbar spinal cord injury
Winklhofer et al, ¹¹ 2013	III	This paper provides evidence that magnetic resonance imaging frequently changes classification of acute traumatic thoracolumbar spine injuries