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

**Sponsored by:** Congress of Neurological Surgeons (CNS) 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).

James S. Harrop, MD,[1] John H. Chi, MD, MPH,[2] Paul A. Anderson, MD,[3] Paul M. Arnold, MD,[4] Andrew T. Dailey, MD,[5] Sanjay S. Dhall, MD,[6] Kurt M. Eichholz, MD,[7] Daniel J. Hoh, MD,[8] Sheeraz Qureshi, MD, MBA,[9] Craig H. Rabb, MD,[10] P. B. Raksin, MD,[11] Michael G. Kaiser, MD,[12] and John E. O'Toole, MD, MS[13]

1. Departments of Neurological Surgery and Orthopedic Surgery, Thomas Jefferson University, Philadelphia, Pennsylvania

2. Department of Neurosurgery, Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts

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, University of Utah, Salt Lake City, Utah

6. Department of Neurological Surgery, University of California, San Francisco, San Francisco, California

7. St. Louis Minimally Invasive Spine Center, St. Louis, Missouri

8. Lillian S. Wells Department of Neurological Surgery, University of Florida, Gainesville, Florida

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

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:***

James S. Harrop, MD

Professor, Departments of Neurological Surgery and Orthopedic Surgery

Director, Division of Spine and Peripheral Nerve Surgery

Neurosurgery Director of Delaware Valley SCI Center

Thomas Jefferson University

909 Walnut Street – Third Floor

Philadelphia, PA 19107

Email: James.Harrop@jefferson.edu

***Keywords:*** Thoracic, thoracolumbar, lumbar, neurologic examination, SCI

***Abbreviations***

AbH – Abductor hallucis

AIS – American Spinal Injury Association Impairment Scale

ASIA – American Spinal Injury Association

FIM – Functional independence measure

SCI – Spinal cord injury

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

## ABSTRACT

**Background:** Traumatic thoracic and lumbar fractures with or without neurologic deficits are less common injuries that typically have been included with all traumatic spine fractures due to their lower prevalence. However, these injuries have unique features in terms of their mechanism of injury, recovery, and neurologic outcomes due to the presence of both upper and lower motor injuries.

**Objective:** To identify neurologic signs and assessment tools that aid in the evaluation and treatment of patients with traumatic thoracic and lumbar fractures.

**Methods:** The guidelines task force initiated a systematic review of the literature relevant to the diagnosis and treatment of patients with thoracolumbar trauma. The National Library of Medicine PubMed database and the Cochrane Library were searched for the period from January 1, 1946 to March 31, 2015, using MeSH subject headings and related keywords.

**Results:** A total of 1195 abstracts were identified. Task force members reviewed all abstracts yielded from the literature search and identified 79 full-text articles for review. Of these, 66 were rejected for not meeting inclusion criteria or for being off topic. Thirteen articles were selected for inclusion in this systematic review. Three articles addressed the question of which neurological assessment tools have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures, and 10 articles addressed the question of whether there are any clinical findings in patients with thoracic and lumbar fractures that can assist in predicting clinical outcomes.

**Conclusion:** Numerous neurologic assessment scales (Functional Independence Measure, Sunnybrook Cord Injury Scale, and Frankel Scale for Spinal Cord Injury) have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures. Unfortunately, other contemporaneous measurement scales (i.e., American Spinal Injury Association Impairment Scale) have not been specifically studied in patients with thoracic and lumbar fractures. However, entry American Spinal Injury Association Impairment Scale grade, sacral sensation, ankle spasticity, urethral and rectal

sphincter function, and abductor hallucis (AbH) motor function can be used to predict neurologic function and outcome in patients with thoracic and lumbar fractures.

## RECOMMENDATIONS

### *Question 1*

Which neurological assessment tools have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures (i.e., do these instruments provide consistent information between different care providers)?

### *Recommendation 1*

Numerous neurologic assessment scales (Functional Independence Measure, Sunnybrook Cord Injury Scale, and Frankel Scale for Spinal Cord Injury) have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures. Unfortunately, other contemporaneous measurement scales (i.e., American Spinal Cord Injury Association Impairment Scale) have not been specifically studied in patients with thoracic and lumbar fractures.

*Strength of Recommendation: Grade C*

### *Question 2*

Are there any clinical findings (e.g., presenting neurological grade/function) in patients with thoracic and lumbar fractures that can assist in predicting clinical outcomes?

### *Recommendation 2*

Entry American Spinal Injury Association Impairment Scale grade, sacral sensation, ankle spasticity, urethral and rectal sphincter function, and AbH motor function can be

used to predict neurologic function and outcome in patients with thoracic and lumbar fractures (Table I).

*Strength of Recommendation: Grade B*

## INTRODUCTION

### *Goals and Rationale*

This clinical guideline has been 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. The surgical management of these patients often takes place under a variety of circumstances and by various clinicians. This guideline has been 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.

Spine fractures can have a devastating effect on patients, particularly because they typically occur in younger populations and often are associated with neurologic injuries. Each patient should be treated as an individual due to the variability and heterogeneity of the patient's mechanism of injury, associated injuries, and accompanying neurologic deficits. These guidelines are centered on thoracic, thoracolumbar, and lumbar fractures. However, variability exists even for these fractures due to the unique biomechanics based on fracture location, association with surrounding anatomical structures, and an individual patient's demographics. Overall patient outcomes are based on numerous factors; however, the patient's neurologic status will have a significant impact on their

prognosis and quality of life. Therefore, this clinical practice guideline focuses on the literature regarding neurologic assessment tools for thoracic and lumbar fractures and sought to evaluate the literature with respect to the following question: which neurological assessment tools have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures (i.e., do these instruments provide consistent information between different care providers)?

In addition, because prognosis is associated with neurologic injuries and deficits, the literature was also reviewed to identify clinical findings that aid in determining the prognosis of neurologic recovery, seeking to address the following question: are there any clinical findings (e.g., presenting neurological grade/function) in patients with thoracic and lumbar fractures that can assist in predicting clinical 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 injury. These guidelines were developed for educational purposes to assist practitioners in their clinical decision-making processes. Additional information about the methods used in this systematic review can be found 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), using 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 1195 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, and/or III literature was available to answer specific questions, the task force did not review level IV studies.

The task force selected 79 full-text articles for review. Of these, 66 were rejected for not meeting inclusion criteria or for being off topic. Thirteen articles 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 do not meet the following criteria were, for the purposes of this evidence-based clinical practice guideline, 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 ≥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 until March 31, 2015;
- 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 evidence-based guideline development methodology. The North American Spine Society methodology 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 the 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 prior to 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.

*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.

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. For 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.”[1] In addition, the task force will confirm 5 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*

Which neurological assessment tools have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures (i.e., do these instruments provide consistent information between different care providers)?

### *Recommendation 1*

Numerous neurologic assessment scales (Functional Independence Measure, Sunnybrook Cord Injury Scale, and Frankel Scale for Spinal Cord Injury) have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures. Unfortunately, other contemporaneous measurement scales (i.e., the American Spinal Injury Association Impairment Scale) have not been specifically studied in patients with thoracic and lumbar injuries. *Strength of Recommendation: Grade C.*

There has been a long evolution of classification and categorization of spinal cord injuries using the neurologic examination as an assessment tool. The origins from which

these assessments appear to have evolved are the Frankel classification scheme. In 1969, Frankel et al[2] reviewed 682 traumatic spinal cord injured (SCI) patients who were treated over a 19-year period. Each patient was categorized based on the degree of neurologic deficit, using a severity grade based on letter grades, from A to E, with A representing a complete neurological deficit and E representing no functional deficit.[2] In 1978, Bracken et al[3] further expanded on Frankel et al's scale[2] by including a 5-point motor scale and 7-point sensory scale tests; this was done in order to have more descriptive categorizations of the injuries and aid with research of SCI patients such that smaller changes in the neurologic examinations could be more accurately detailed. In 1982, the American Spinal Injury Association (ASIA) developed an international standard for neurological and functional classification of spinal cord injury.[4] The ASIA examination consisted of a detailed motor examination on a 0 to 5 scale with 10 motor groups that represented discrete spinal cord segments (e.g., triceps motor strength represented C7 neurologic distribution). However, this classification system did not include a detailed sensory component. The ASIA examination has undergone several revisions and has evolved to employ motor, sensory, ASIA impairment score, and a functional impairment measure (FIM). However, one deficiency in these assessment tools has been their heterogeneity and inclusion of cervical, thoracic, lumbar, and sacral injuries.

These guidelines focus on thoracic and lumbar fractures, and as such, the inclusion criteria require that 80% of the patients had to have thoracic or lumbar injuries. Thus, with these restrictions, only three articles were identified that focused on the neurologic assessment of thoracic and lumbar fractures patients. Davis et al[5] performed a prospective

review of 43 thoracic and lumbar patients and evaluated the reliability of the Frankel and Sunnybrook scales. The patients were assessed by 3 physical therapists and scores were then recorded in a database, and were subsequently compared. The Pierson correlation coefficients were noted to be high (0.71-0.91). Although the interrater reliability was high with both scales, ranging from 94% to100%, there was better agreement in terms of interrater reliability with the Frankel scale over the Sunnybrook scale.[5] However, both scales were deemed insensitive in that significant recovery in a patient's motor, sensory, bladder, or walking functions occurred without any change in their scale. In addition, a heterogeneous patient population was studied, with the inclusion of thoracolumbar with lumbar fractures (included L3-5).[5] This prospective diagnostic study was downgraded to level III due to a lack of universal applied reference "gold" standard, heterogeneous cohort, and failure to report sensitivity and specificity.

The other 2 studies investigated the FIM as an assessment tool and also concentrated on thoracic and lumbar fractures. Beck et al[6] retrospectively reviewed 56 traumatic thoracic and lumbar patients and concluded that a thoracic SCI patient's disposition could be based on the level of spine injury and the completeness of SCI alone using the FIM assessment tool. This was a retrospective diagnostic study but was downgraded to level IV due to a lack of universal applied reference "gold" standard, poor homogeneity of subjects, small sample size, nonconsecutive, and failure to use validated outcome measures.[6] Barbetta et al[7] reported the FIM was valid and responsive for thoracic and lumbar fractures, and neurologic injuries in a large series of 218 Brazilian individuals with spinal cord injury.[5] This retrospective diagnostic study was downgraded to level III

due to a lack of universal applied reference “gold” standard and a failure to report correlation coefficients.

***Question 2***

Are there any clinical findings (e.g., presenting neurological grade/function) in patients with thoracic and lumbar fractures that can assist in predicting clinical outcomes?

***Recommendation 2***

Entry ASIA Impairment Scale grade, sacral sensation, ankle spasticity, urethral and rectal sphincter function, and AbH motor function can be used to predict neurologic function and outcome in patients with thoracic and lumbar fractures (Table I). *Strength of Recommendation: Grade B.*

There are numerous level III and IV retrospect prognostic studies on thoracic and lumbar fractures and their association with neurologic deficits and assessment techniques These studies show that patients with more severe neurologic injuries had worse neurologic outcomes in terms of recovery.[8-12]

Benzel et al[13] retrospectively reviewed 105 anteriorly decompressed and fused cases of thoracic and lumbar fractures and noted that none of the 34 patients with complete motor and sensory loss had any return of function. McLain[11] retrospectively reviewed the return to work status at 5-year follow-up after injury and used the Frankel grade on 70 thoracic, thoracolumbar, and lumbar spine fractures that had a variety of operative treatments. He also reported that the patient’s neurologic injury had a greater impact on functional

outcome over any other variable.[11] Dobran et al[9] further noted that the neurologic examination or admission ASIA grade of patients undergoing a posterior approach for thoracolumbar fractures was the strongest predictive factor of neurological improvement in univariate analysis ($p$ = .0005). These authors in an additional multivariate analysis reported that preoperative neurological status ($p$ = .0491) and the fracture type ($p$ = .049) had a positive predictive value on neurologic outcome. These 3 retrospective prognostic studies were downgraded to level III due to a high degree of variance and heterogeneity of treatment, lack of detail in reporting, and nonconsecutive enrollment.

Harrop et al[10] retrospectively reviewed 94 spine trauma patients and categorized them by the level of injury as thoracic (T4-9), thoracolumbar (T10-T12) and lumbar SCI, and noted that the lumbar or conus injuries had the greatest neurologic recovery as graded by the ASIA classification. The authors not only reported the severity of the injury but also the level of the injury that resulted in neurologic recovery. They attributed the improved recovery to the higher concentration of lower motor neurons and the ability of the neurons to develop "root escape."[10] This was downgraded to a level III prognostic study due to its retrospective nature, lack of follow-up (95/150 patients), and being nonconsecutive. Kingwell et al[14] also illustrated that the anatomic level of injury based on neurologic examination is a better predictor of recovery than the MRI fracture location. It was downgraded to level II due to the retrospective nature of the study. The only study to use validated outcome measures (the Short-Form-36, Oswestry Disability Index, and Prolo Economic Scale outcome instruments were completed at a minimum follow-up of 12 months) was performed in a thoracic (T1-10) population with and without neurologic

deficits. Schouten et al[12] noted in these 126 cases, patients with neurologic injuries graded by the ASIA classification had worse outcome measures. All of these 3 studies were also downgraded to level III evidence due to lack of a reference "gold" standard. This prognostic study was downgraded to level III due to the retrospective nature of the study in addition to lack of homogeneity and limited response rate of 39%.

In addition to the larger retrospective series on neurologic injury and outcome, several authors have identified clinical indications and neurologic recovery.[14-18] Calancie et al[15] performed prospective electromyograms (EMGs) and clinical evaluations of 70 incomplete SCI patients with 58 lesions rostral to T10 and T12 with caudal injuries. The authors noted that the AbH was an earlier and more accurate indicator of supraspinal influence and the recovery of neurologic function. This prospective prognostic study was downgraded to level II due to lack of follow-up and consistent of information. Kingwell et al[14] noted that the absence of initial sacral sensation had a negative effect on motor recovery by a factor of 13.2 points. Chen et al[16] also used lower extremity neurologic function in a prospective review of 52 patients with incomplete neurologic injuries from thoracic and lumbar fractures where 50% were epiconus injuries. Sensitivity, specificity, positive predictive value, and negative predictive value of increased knee jerk on detrusor/sphincter spasticity were 27.6%/25%, 100%/100%, 100%/100%, and 48.8%/41.5%, and increased ankle jerk on detrusor/sphincter spasticity were 64.5%/61.8%, 95.2%/100%, 95.2%/100%, and 64.5%/58.1%, respectively. Overall, the authors concluded that in thoracolumbar fracture patients with neurologic deficits, ankle spasticity is highly accurate in predicting neurogenic bladder dysfunction.[16] This

prospective prognostic study was downgraded to level II due to heterogeneity of treatments and patient population, no anatomic confirmation of injury level, and nonconsecutive population.

Schurch[17] prospectively examined 63 patients with thoracolumbar fractures and SCI using the ASIA protocol and urodynamics. Seven patients recovered from their neuropathic voiding disorders, and there was a significant correlation between the reappearance of a voluntary external anal/urethral sphincter contraction and bladder recovery ($p < .0l$). In a later report, Schurch et al[18] noted that in thoracolumbar SCI patients, pinprick sensation in the perineal area is of negative predictive value. Specifically, the absence of pinprick sensation predicts poor bladder recovery.[17] This prospective prognostic study was downgraded to level II due to a high degree of heterogeneity of the patient population.

**Future Research**

This systematic review of the literature identifies several areas for future research and investigation. There have been numerous publications and advancement in the examination of patients with spinal cord injuries. Unfortunately, these studies include all patients with neurologic injuries (cervical, thoracic, and lumbar), and there has not been categorization by anatomic levels. Recent clinical trials and investigations have begun to separate cervical from thoracic and lumbar injuries. This is due to their different mechanism of injury, concurrent injuries, and recovery pattern. Thus, future work will

need to be focused on the assessment and prognostic tools for isolated thoracic and lumbar injuries.

## Conclusions

There is limited research and literature that focuses specifically on thoracic and lumbar fracture patients. Despite these limitations, there are numerous neurologic assessment scales (FIM, Sunnybrook, and Frankel) that have demonstrated internal reliability and validity in the management of patients with thoracic and lumbar fractures. Unfortunately, other contemporaneous measurement scales (ie, ASIA) have not been specifically studied in patients with thoracic and lumbar fractures. However, entry AIS grade, sacral sensation, ankle spasticity, urethral and rectal sphincter function, and AbH motor function can be used to predict neurologic function and outcome in these patients.

### *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. Ransohoff DF, Pignone M, Sox HC. How to decide whether a clinical practice guideline is trustworthy. JAMA 2013;309:139-140.

2. Frankel HL, Hancock DO, Hyslop G, et al. The value of postural reduction in the initial management of closed injuries of the spine with paraplegia and tetraplegia. I. Paraplegia 1969;7:179-192.

3. Bracken MB, Webb Jr SB, Wagner FC. Classification of the severity of acute spinal cord injury: implications for management. Paraplegia 1978;15:319-326.

4. Priebe MM, Waring WP. The interobserver reliability of the revised American Spinal Injury Association standards for neurological classification of spinal injury patients. Am J Phys Med Rehabil 1991;70:268-270.

5. Davis LA, Warren SA, Reid DC, Oberle K, Saboe LA, Grace MG. Incomplete neural deficits in thoracolumbar and lumbar spine fractures. Reliability of Frankel and Sunnybrook scales. Spine (Phila Pa 1976) 1993;18:257-263.

6. Beck LA, Harris MR, Basford J. Factors influencing functional outcome and discharge disposition after thoracic spinal cord injury. SCI Nurs 1999;16:127-132.

7. Barbetta DC, Cassemiro LC, Assis MR. The experience of using the scale of functional independence measure in individuals undergoing spinal cord injury rehabilitation in Brazil. Spinal Cord 2014;52:276-281.

8. Benzel EC, Larson SJ. Functional recovery after decompressive operation for thoracic and lumbar spine fractures. Neurosurgery 1986;19:772-778.

9. Dobran M, Iacoangeli M, Di Somma LG, et al. Neurological outcome in a series of 58 patients operated for traumatic thoracolumbar spinal cord injuries. Surg Neurol Int 2014;5(suppl 7):S329-S332.

10. Harrop JS, Naroji S, Maltenfort MG, et al. Neurologic improvement after thoracic, thoracolumbar, and lumbar spinal cord (conus medullaris) injuries. Spine (Phila Pa 1976) 2011;36:21-25.

11. McLain RF. Functional outcomes after surgery for spinal fractures: return to work and activity. Spine (Phila Pa 1976) 2004;29:470-477.

12. Schouten R, Keynan O, Lee RS, et al. Health-related quality-of-life outcomes after thoracic (T1-T10) fractures. Spine J 2014;14:1635-1642.

13. Benzel EC, Larson SJ. Recovery of nerve root function after complete quadriplegia from cervical spine fractures. Neurosurgery 1986;19:809-812.

14. Kingwell SP, Noonan VK, Fisher CG, et al. Relationship of neural axis level of injury to motor recovery and health-related quality of life in patients with a thoracolumbar spinal injury. J Bone Joint Surg Am 2010;92:1591-1599.

15. Calancie B, Molano MR, Broton JG. Abductor hallucis for monitoring lower-limb recovery after spinal cord injury in man. Spinal Cord 2004;42:573-580.

16. Chen SL, Huang YH, Wei TY, Huang KM, Ho SH, Bih LI. Motor and bladder dysfunctions in patients with vertebral fractures at the thoracolumbar junction. Eur Spine J 2012;21:844-849.

17. Schurch B. The predictive value of plantar flexion of the toes in the assessment of neuropathic voiding disorders in patients with spine lesions at the thoracolumbar level. Arch Phys Med Rehabil 1999;80:681-686.

18. Schurch B, Schmid DM, Kaegi K. Value of sensory examination in predicting bladder function in patients with T12-L1 fractures and spinal cord injury. Arch Phys Med Rehabil 2003;84:83-89.

**Table I. Assessment Tools and Outcomes**

| Neurological Assessments Tool | Outcomes (Correlated) |
|---|---|
| Worse ASIA Impairment Scale grade entry score | Worse neurologic outcomes and greatest neurologic deficits |
| Absence of sacral sensation on initial evaluation | Worse recovery of neurologic function |
| Absence of sacral pinprick sensation | Poor prognosis for bladder recovery |
| Identification of ankle spasticity | Presence of neurogenic bladder dysfunction |
| Reappearance of urethral and rectal sphincter presence | Bladder recovery |
| Presence of electromyographic evidence of abductor hallucis motor function | The earliest and most accurate predictor of neurologic recovery |

ASIA, American Spinal Injury Association.

## Appendix I. Literature Searches

### *Search Strategies*

### PubMed

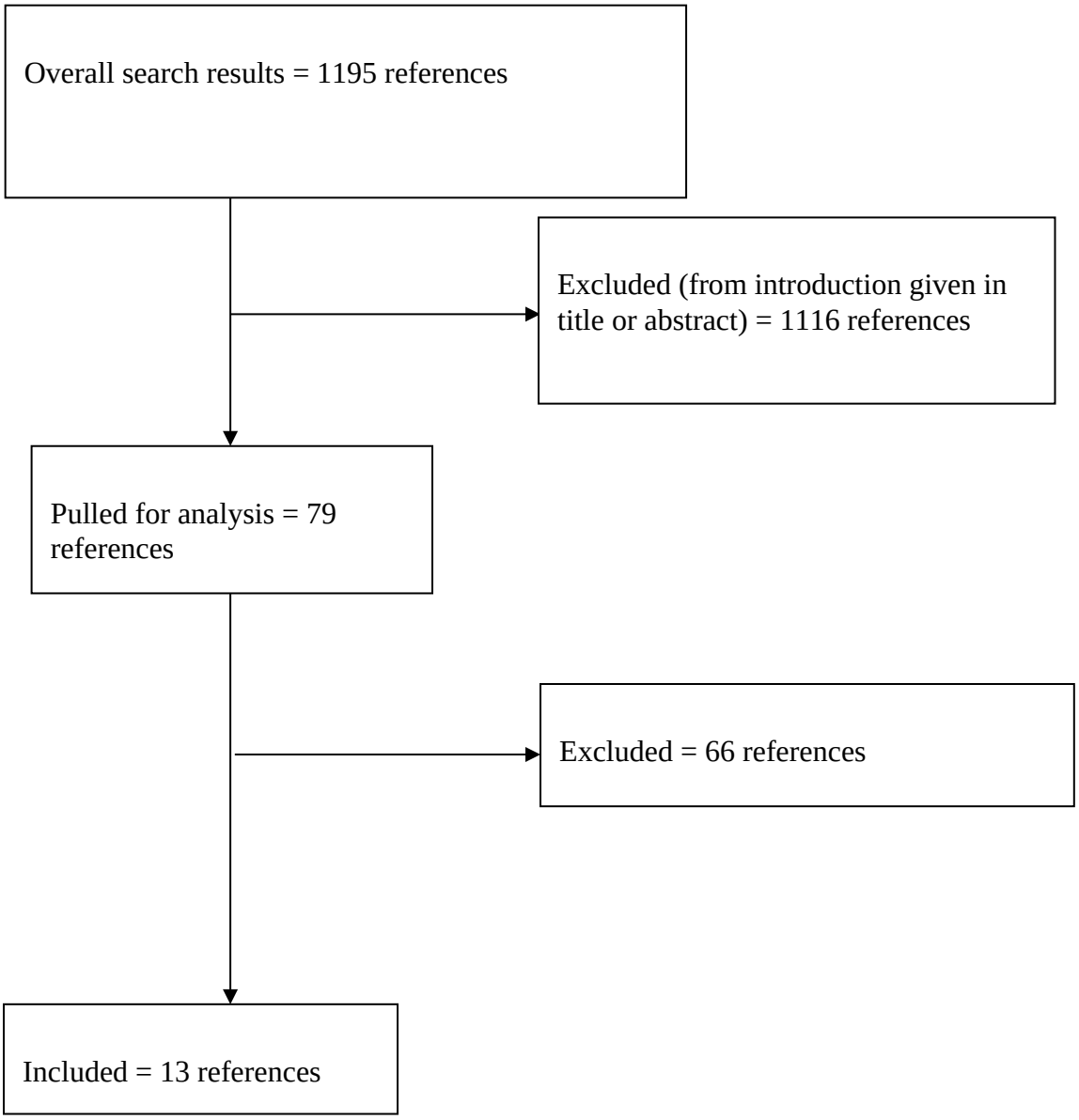
1. Lumbar vertebrae [MeSH] OR Thoracic vertebrae [MeSH]
2. Spinal Injuries [MeSH] OR Spinal Cord Injuries [MeSH]
3. #1 AND #2
4. Thoracolumbar [TIAB] OR thoraco-lumbar [TIAB] OR thoraco lumbar [TIAB] OR burst [Title]
5. Injur* [TIAB] OR trauma* [TIAB] OR fractur* [TIAB] OR dislocation* [TIAB]
6. #4 AND #5
7. Lumbar vertebrae/injuries [MeSH] OR Thoracic vertebrae/injuries [MeSH] (3150 results)
8. #3 OR #6 OR #7
9. Trauma Severity Indices [Mesh] OR Neurologic examination [MeSH]
10. (Neurologic* [TIAB] OR Motor [TIAB] OR Sensory [TIAB]) AND (assessment*[Title/Abstract] OR examination* [TIAB] OR exam [TIAB] OR exams [TIAB] OR test [TIAB] OR tests [TIAB] OR testing [TIAB] OR evaluat* [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* 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. osteoporo* 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 | • High-quality randomized trial with statistically significant difference or no statistically significant difference but narrow confidenceintervals<br>• Systematic review[b] of level I RCTs (and study results were homogenous[c]) | • High-quality prospective study[d] (all patients were enrolled at the same point in their disease with ≥80% follow-up of enrolled patients)<br>• Systematic review[b] of level I studies | • Testing of previously developed diagnostic criteria on consecutive patients (with universally applied reference "gold" standard)<br>• Systematic review[b] of level I studies | • Sensible costs and alternatives; values obtained from many studies; with multiway sensitivity analyses<br>• Systematic review[b] of level I studies |

| | | | | |
|---|---|---|---|---|
| Level II | • Lesser quality RCT (e.g., ≤80% follow-up, no blinding, or improper randomization)<br>• Prospective[d] comparative study[e]<br>• Systematic review[b] of level II studies or level I studies with inconsistent results | • Retrospective[f] study<br>• Untreated controls from an RCT<br>• Lesser quality prospective study (e.g., patients enrolled at different points in their disease or ≤80% follow-up)<br>• Systematic review[b] of level II studies | • Development of diagnostic criteria on consecutive patients (with universally applied reference "gold" standard)<br>• Systematic review[b] of level II studies | • Sensible costs and alternatives; values obtained from limited studies; with multiway sensitivity analyses<br>• Systematic review[b] of level II studies |
| Level III | • Case control study[g]<br>• Retrospective[f] comparative study[e]<br>• Systematic review[b] of level III studies | • Case control study[g] | • Study of non consecutive patients; without consistently applied reference "gold" standard<br>• Systematic review[b] of level III studies | • Analyses based on limited alternatives and costs; and poor estimates<br>• Systematic review[b] of level III studies |
| Level IV | Case series[h] | Case series | • Case-control study<br>• Poor reference standard | • Analyses with no sensitivity analyses |

RCT, Randomized controlled trial.

[a]A complete assessment of quality of individual studies requires critical appraisal of all aspects of the study design.

[b]A combination of results from ≥2 previous studies.

[c]Studies provided consistent results.

[d]Study was started before the first patient enrolled.

[e]Patients treated one way (e.g., instrumented arthrodesis) compared with a group of patients treated in another way (e.g., unsintrumented arthrodesis) at the same institution.

[f]The study was started after the first patient enrolled.

[g]Patients 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).

[h]Patients 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] |

[a]Note 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;
    - <80% of patient follow-up;
    - Failure to utilize validated outcomes instrument;
    - No statistical analysis of results;
    - Cross over rate between treatment groups of >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).

3. Methodology of diagnostic studies reviewed for following deficiencies:
    - Failure to determine specificity and sensitivity;
    - Failure to determine inter- and intraobserver 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 |
|---|---|---|
| Barbetta et al,[7] 2014 | III | This paper provides evidence that FIM is valid assessment tool for thoracic and lumbar neurologic injuries |
| Beck et al,[6] 1999 | IV | This paper provides evidence that completeness of neurologic injury and level of injury are sufficient to determine disposition in thoracic SCI population |
| Benzel et al,[13] 1986 | III | This paper provides evidence that worse neurologic injury lead to worst prognosis |
| Calancie et al,[15] 2004 | II | This paper provides evidence that prognostic assessment of intrinsic foot muscles via electromyogram may be important metric |
| Chen et al,[16] 2012 | II | This paper provides evidence that ankle spasticity in chronic thoracolumbar fracture patient suggests upper motor urinary system dysfunction |
| Davis et al,[5] 1993 | II | This paper provides evidence that Frankel and Sunnybrook neurologic assessment scales in the thoracic and lumbar patients with neurologic deficits were reliable and both had high intra- and interrater reliabilities |
| Dobran et al,[9] 2014 | III | This paper provides evidence that posterior operated thoracolumbar fracture patients' neurologic examination or ASIA grade at admission ($p$ = .0005) was the strongest predictive factor of neurological improvement in univariate analysis |
| Harrop et al,[10] 2011 | III | This paper provides evidence that the lumbar or conus injuries had the greatest neurologic recovery as graded by the ASIA classification |
| Kingwell et al,[14] 2010 | II | This paper provides evidence that anatomic level of injury by MRI and ASIA exam correlates better than skeletal level after thoracolumbar injury |
| McLain,[11] 2004 | III | This paper provides evidence that: neurologic injury at time of injury was greatest predictor to ability to return to work |
| Schouten et al,[12] 2014 | III | This paper provides evidence that thoracic injuries T2-10 had overall good functional outcomes compared to general population. Those patients with significant neurologic injuries had worse outcomes |

| Author, Year | Level of Evidence | Task Force Conclusions Relative to Question and Rationale for Evidence Grading |
|---|---|---|
| Schurch,[17] 1999 | II | This paper provides evidence that there was no differentiation between motor function at toes and bladder dysfunction after SCI |
| Schurch et al,[18] 2003 | III | This paper provides evidence that sensory examination alone does not predict voiding function |

ASIA, American Spinal Injury Association; FIM, functional independence measure; MRI, magnetic resonance imaging; SCI, spinal cord injury.