Review Article | **Open Access**

Thoracolumbar Injury Severity Scoring Systems: A Review and Rationale for a New System Based on the AOSpine Thoracolumbar Injury Classification System

Robyn Rubenstein[1], John D Koerner[1,2]*, David Oh[9], Chris Kepler[1,2], Frank Kandziora[3], Rajasekaran Shanmuganathan[4], Marcel Dvorak[5], Bizhan Aarabi[6], Luiz Vialle[7], Cumhur Oner[8] and Alexander R Vaccaro[1,2]

[1]Thomas Jefferson University, Philadelphia PA, USA

[2]Rothman Institute, Philadelphia, USA

[3]Center for Spine Surgery and Neurotraumatology, BGU Hospital, Frankfurt, Germany

[4]Ganga Hospital, Coimbatore, Tamil Nadu, India

[5]Vancouver General Hospital, Vancouver, British Columbia, Canada

[6]University of Maryland School of Medicine, Baltimore, USA

[7]Catholic University, Curitiba, Brazil

[8]University Medical Center, Utrecht, The Netherlands

[9]Drexel University College of Medicine, Philadelphia PA, USA

*Corresponding author: John D Koerner, Thomas Jefferson University, Philadelphia, USA, Tel: 215-955-9773; E-mail: johndkoerner@gmail.com

Abstract

Numerous classification systems have been developed over the years to describe thoracolumbar injuries, each with their own benefits and limitations. None of these systems have been accepted however, as a universal, comprehensive system to classify these injuries. The AOSpine Thoracolumbar Injury Classification System has recently been developed in order to overcome some of the limitations of previous systems. An injury severity scoring system based on this system would be beneficial to clinicians when treating these complex injuries. This paper will review thoracolumbar injury classification systems, and describe the need for a new severity scoring system based on the AOSpine Thoracolumbar Injury Classification System.

Introduction

The systematic classification of thoracolumbar injuries has been altered, expanded, and repeatedly critiqued since Böhler's classification in 1929 [1]. Various classification systems have been proposed over the last 80 years that emphasize different aspects of thoracolumbar trauma, diagnosis, and prognosis. In 1949, Nicoll classified spinal injuries by dividing them into stable and unstable injuries, which he argued was essential in choosing the appropriate treatment [2]. Holdsworth reemphasized the importance of distinguishing between stable and unstable spinal fractures while stating that there were other factors involved as well, including the integrity of the posterior ligamentous complex (PLC) [3]. He divided spinal injuries into five groups which were based on the mechanism of the injury [4]. Denis then introduced the three-column model of the spine describing an anterior, middle, and posterior column, all of which are susceptible to various types of injuries [5]. Denis' classification system has been critiqued as being too simple and failing to recognize a variety of potential fractures and injuries [6]. McAfee et al. elaborated on Denis' three column spine system and emphasized the importance of computed tomography in revealing essential aspects of an injury [7]. Ferguson and Allen focused on the mechanism of injury as the basis of their classification system [8].

While these systems classify and compartmentalize the various types of thoracolumbar injuries or present the injuries in a hierarchical manner ordered by severity, none of these systems have been accepted as a universal, comprehensive system [6]. The tendency is to rely on the Magerl or AO Comprehensive Classification in Europe, while in North America variations of the Denis, McAfee, and TLICS systems are preferred, which has perpetuated the difficulty of adopting a comprehensive system. In spite of many advances in imaging and surgical techniques for treating fractures, there have been several inherent limitations that have plagued the development of spinal injury classification systems [9]. Despite the many systems, none have adequately described injury severity, pathogenesis, and biomechanical inciting forces while addressing all clinical, neurological and radiological characteristics which are important for making treatment decisions [10].

This lack of acceptance of previous systems is in part due to the difficulty of using these systems in a clinical setting as well as the lack of reproducibility of many of these classifications [6]. For example, Wood et al. reported that both the Denis and Magerl systems indicated only moderate reliability and repeatability, with an average kappa coefficient of 0.475 for the Magerl system (in regards to assigning one of the three types of injuries) and 0.606 for Denis classification system (in regards to assigning one of the four fracture types) [11]. Furthermore, many of these systems are either too complex or are oversimplified, creating confusion, limiting their clinical usefulness, and hindering their ability to improve communication between clinicians and to educate residents and fellows [6]. Clinicians and researchers are in need of a system that is simple yet simultaneously all-encompassing that can promote effective and progressive

communication in the field of spinal trauma and treatment. Achieving reliability has been a major challenge in creating classifications systems, as the algorithmic process must also compress available information into reproducible categories without loss of information content. This compression progresses towards two predictable pitfalls: 1) either there is a loss of information content in favor of simplicity and higher reproducibility, or 2) there is a loss of simplicity and reproducibility in favor of higher informational content [6]. Thus, the primary concern for current classification of thoracolumbar injuries has been limitation by either excessive complexity or lack of inclusiveness [10,12]. Furthermore, many of these classification systems serve only to describe the types of injuries, but do not direct clinicians to appropriate treatment [6].

This paper will discuss two commonly referenced and discussed classification systems, Magerl and TLICS, as well as a system which incorporates elements of both systems, the AOSpine Thoracolumbar Injury Classification System. The effectiveness of the AOSpine Classification System will be discussed as well as the need for this system to incorporate a severity scoring system similar to TLICS in order to guide treatment. This review will also attempt to identify some of the challenges in creating a scoring system based on the critiques of former systems.

Magerl Classification System

The Magerl classification system, created in 1994, is arguably the most detailed, complex, and systematic spinal trauma classification system [13]. The system, based on the 3-3-3 scheme of the AO fracture classification, distinguishes three main types by the injury morphology: type A (vertebral body compression), type B (distraction), and type C (rotation due to axial torque) [14]. Each type has three groups, and each group consists of three subgroups, which can have further specifications as well. The degree of severity of each injury, expressed in terms of instability and comminution, is stratified by its placement in the classification system, with increasing severity going from type A injuries to type C injuries. The fundamental injury patterns discussed in the Magerl classification system were diagnosable by radiographs and CT scan [14].

Although this system is comprehensive, many argue that it is probably overly complex, creating unnecessary confusion for clinicians and researchers alike [6,13]. The Magerl system discusses 17 different type A injuries, 15 different type B injuries, and 23 different type C injuries. This complexity stymied the effectiveness and clinical usage of this scheme. Furthermore, Magerl et al. noted that type C injuries are often superimposed on type A injuries, enhancing the complexity of the system [14]. Additionally, there is insufficient discussion of treatments for the various types of injuries. While the study presents 55 different types of injuries, there are only three small sections in the original article which describe treatment of injuries. This discussion of treatment lacked detail and failed to provide guidance for clinical decision-making [6]. Additionally, although Magerl et al. acknowledged the correlation between the level of injury severity and the frequency of neurological deficit, the system failed to offer concrete information about how the increased likelihood of neurological injury and the existence of additional clinical factors and comorbidities should impact the choice of treatment [13]. Another potential issue with this system is the confusion caused by transitional type injuries that span several categories in the classification. For example, a type A injury can become type B when the degree of flexion exceeds the estimated point beyond which the PLC will fail [14]. In addition, users had difficulty in identifying specific types of injuries, as many posterior distraction injuries could be mistaken for type A injuries [14]. Although the Magerl system was designed to be a morphological model, ultimately it was overly descriptive and did not help in guiding clinical decision-making.

TLICS

Vaccaro et al. created the TLICS with the intention of overcoming the failure of other classification systems to provide concrete direction for optimal treatment of thoracolumbar injuries [6]. TLICS aims to direct clinicians towards an appropriate path of treatment with the proposal of a severity scoring scale [6]. TLICS is based on three major features of thoracolumbar traumatic injuries: 1) morphology of the injury (based on imaging studies), 2) the integrity of the PLC, and 3) neurologic status (Table 1) [6]. In contrast to the complex nature of Magerl, TLICS simplifies the defined morphologies and types of injuries, allowing combinations of morphologies to denote injuries of greater severity rather than listing all potential types [6].

The most unique and progressive aspect of TLICS is its injury severity scoring system, which assigns values dependent on whether the injury morphology is compression (1 point), compression with burst component (2 points), translational/rotational (3 points), or distraction (4 points). In the case of multiple injuries, the clinician should assign a score for the most severe (highest scoring) injury. Furthermore, the neurologic status of the patient is also given a score depending on the urgency of need for surgical decompression. Neurological deficit scoring is as follows: neurologically intact (0 points), nerve root injury (2 points), complete spinal cord injury (2 points), incomplete sensory or motor spinal cord injury (3 points), cauda equina injury (3 points). The composite score assists in determining if surgical intervention is encouraged or not. A score of 3 or below represents a nonoperative injury, and a score of 5 or above suggests that surgical intervention should be considered. A score of 4 is indeterminate and the clinician must use his/her experience and judgment to decide if the injury will be treated conservatively or surgically [6].

Furthermore, TLICS offers a system of suggested surgical approaches depending on the neurologic status of the patient and the integrity of the PLC. For patients with an incomplete neurologic injury, an anterior approach is necessary if there is neural compression from anterior structures. If there is a PLC injury, a posterior approach is generally necessary. When both of these scenarios are present, an anterior posterior approach is required [6]. Additionally, Vaccaro et al. noted the significance of taking into account, along with the injury severity score, the local clinical considerations, remote comorbidities, and systemic considerations while using TLICS [6]. These various clinical modifiers can cause a nonsurgical injury to require surgical intervention or vice versa [6].

Although a TLICS score of 4 is indeterminate, and may be criticized for its inability to direct treatment, a system that in no way requires clinical judgment is unrealistic; it is inevitable that there will be situations where anecdotal experience must assist in guiding treatment. Furthermore, systematic guidance towards surgical decisions cannot replace a surgeon's past experiences, and there are often external factors making each patient's situation more complex, not only benefitting from but at times requiring a certain level of experience to assist in providing the optimal treatment [6]. Therefore, TLICS is a guide towards the appropriate treatment, but not a

replacement for a surgeon's intuition [6]. There has been criticism over the fact that TLICS may not be applicable in every and all cultures and regions, especially in areas where MRI is not available [13]. Some argued that the severity scoring system did not necessarily reflect universal surgical practices or the most sensible and realistic method of treatment [13].

The reliability of TLICS has been evaluated by numerous groups since its introduction [15-19]. While there are many advantages in using the TLICS in determining treatment, studies have discovered certain injuries that are not well identified with this system. In a recent study by Moore et al. [20], low lumbar burst fractures were evaluated by 15 fellowship-trained spine surgeons. The final TLICS score had a 28% agreement with a kappa of 0.245. The reliability improved significantly for L3 injuries compared to L4 and L5 injuries, suggesting there may be inherent differences in injury identification with lower lumbar burst fractures [20]. Some studies have found that determining the integrity of the PLC to be the most difficult aspect of TLICS [16]. Others have reported cases of progressive kyphotic deformity that develops after conservative treatment of comminuted burst fractures, and suggest that the recommendation for nonoperative management of neurologically intact burst fractures is a pitfall of the system [21]. However, another study reported on the clinical outcomes of patients treated according to TLICS, and found no neurological worsening in sixty-five patients treated according to the system [22]. The potential pitfalls and benefits of TLICS provided valuable guidance in the development of a new numerical scoring system based on the AOSpine classification system.

AOSpine

The AOSpine system was created because previous systems had not found the "ideal mix between simplicity and comprehensiveness" required to establish a universal system [13]. The goal of the system was for it to be clinically useful, accepted universally, and easy to use. The AOSpine Thoracolumbar Injury Classification system, compositely developed by an international team of clinicians and researchers, integrates aspects of both the Magerl system and TLICS [13].

The AOSpine classification system takes into account: 1) morphologic classification of the injury, 2) grading of neurological status, and 3) acknowledgement and incorporation of significant clinical patient-specific modifiers and comorbidities. The revised AOSpine system uses the information provided by the three main injury categories from the original Magerl AO concept, namely: A) compression B) tension band and C) displacement type injuries [23]. Type A injuries were divided into subtypes (A0, A1, A2, A3, A4) and Type B were divided as well (B1, B2, B3), while all Type C injuries were not subdivided [23].

The AOSpine system requires that multilevel injuries be classified individually and then listed in order from most severe to least severe [13]. While the subtypes resemble the Magerl system, they lack the complexity that was previously the target of criticism and represent distinct morphologic injury patterns rather than a spectrum of less stable similar injuries. Along with the higher reproducibility and simplicity, the reliability and accuracy of injury evaluation under the new AOSpine classification parameters allow the system to take into account the variability and diversity of spinal cord injuries without loss of information content [23]. Thus, the recent AOSpine injury classification system provides a revised scheme that reduces the complexity of previous systems, while still maintaining clinically relevant classification reliability.

Injury Morphology		
Type	**Qualifiers**	**Points**
Compression		1
	Burst	1
Translational/rotational		3
Distraction		4
Integrity of Posterior Ligamentous Complex		
Intact		0
Suspected/Indeterminate		2
Injured		3
Neurologic Status		
Involvement	Qualifiers	
Intact		0
Nerve Root		2
Cord, conus medullaris	Complete	2
	Incomplete	3
Cauda Equina		3
Injury Morphology		
Type	Qualifiers	Points
Compression		1
	Burst	1
Translational/rotational		3
Distraction		4
Integrity of Posterior Ligamentous Complex		
Intact		0
Suspected/Indeterminate		2
Injured		3
Neurologic Status		
Involvement	Qualifiers	
Intact		0
Nerve Root		2
Cord, conus medullaris	Complete	2
	Incomplete	3
Cauda Equina		3

Table 1: Thoracolumbar Injury Classification and Severity Score (TLICS) [6].

Furthermore, the system provides for consideration of neurological deficits in assessing overall injury severity. The neurological status can be designated into any of 6 categories: N0 (neurologically intact), N1 (transient neurological deficit that is no longer present), N2 (symptoms or signs of radiculopathy), N3 (incomplete spinal cord or cauda equina injury), N4 (complete spinal cord or cauda equina injury), and NX (inconclusive due to inability to complete neurological examination). Another valuable aspect of the AOSpine system is that it accounts for patient-specific modifiers and comorbidities, similarly to but more emphatically than TLICS. These modifiers should be considered on an as-needed basis to assist the clinician in determining the pathway of treatment. In the system, M1 represents fractures with an indeterminate injury to the PLC on either clinical examination or imaging studies, while M2 represents a patient-specific comorbidity that may either encourage or hinder a potential surgical treatment [13]. The relegation of indeterminate PLC injury to a modifier rather than as a critical element of the system was intended to deemphasize PLC evaluation, given the relatively low reliability of MRI in this regard [24,25].

The AOSpine classification system demonstrated improved reliability for both interobserver agreement and intraobserver reproducibility. Other systems, like the Magerl classification system, failed to reach this significant reliability [13]. For example, while the AOSpine classification system reported a K coefficient of 0.72 for the identification of the main injury types, the Magerl system reported κ coefficients of only 0.33 and 0.62 [13]. High interobserver and intraobserver reliability is crucial in creating a useful and consistent classification system.

Ultimately, the AOSpine classification system is an amalgamation of some of the most effective and useful aspects of both the Magerl system and TLICS. It proposes a more comprehensive and effective approach to classifying spinal injuries. Similar to the Magerl system, the AOSpine classification system includes careful morphologic description of injuries, with each successive injury representing increased instability and likelihood of need for stabilization [13]. AOSpine's similarity to TLICS is appreciated by noting the inclusion of neurologic status, and considering the significance of patient-specific comorbidities and modifiers. Finally, the AOSpine system does not obligate the use of MRI and relies mostly on CT scan for primary classification, a feature which may increase its utility in the developing world where access to MRI is often limited.

Need for a Severity Scoring System to Guide Treatment

One of the principle challenges in creating a universally accepted severity scoring system for thoracolumbar trauma is the incorporation of various regional and worldwide differences in existing treatment practices. Such differences exist for many reasons, such as availability of resources for diagnostic imaging including MRI, the familiarity surgeons have with certain classifications and how they are taught, the availability and affordability of modern surgical instrumentation, as well as the expectations of patients and surgeons. The clinician's perception of the severity of an injury may be altered by the use of advanced imaging (MRI or CT) and limited access to them may affect the pre-treatment workup and in some cases may change treatment decisions. If the goal of creating a severity scale is to help guide treatment, cultural differences will need to be incorporated to reflect willingness to undergo and perform surgery and financial considerations which may limit or promote operative intervention. For example, certain regions may be more influenced by time missed

from work, and favor an operative intervention over conservative management if the patient returns to work sooner even if long term outcomes are identical. Acceptance of residual deformity may also vary between different cultures.

A clinically useful and comprehensive thoracolumbar trauma classification system should guide treatment decisions. Except for TLICS, previous systems were descriptive schemes and did not offer clinical guidance which took into account the modern diagnostic and therapeutic techniques available [6]. There is a pressing need to create a treatment algorithm based on the recently developed AOSpine classification system. The next step in developing a treatment algorithm is to design a scoring system to stratify injury severity and need for surgical intervention taking into consideration the issues discussed above regarding resource allocation and cultural attitudes toward surgical intervention. Such a system should consider spinal biomechanical stability in short and longer term and neurologic injury or threat thereof. Most injury patterns clearly benefit or do not benefit from surgical stabilization with only a few injuries proving controversial with respect to indications for stabilization. A severity scoring system should give insight into the likely outcome of different treatment methods and may trigger surgical intervention at different thresholds to reflect the clinical and cultural equipoise [6]. Integration of a severity scale, similar to the TLICS, based on the AOSpine system would assist the clinical decision-making process and allow for better communication between researchers and clinicians and clearer and more effective acute injury management by residents, fellows and treating physicians. The criticisms of TLICS should be thoroughly evaluated in order to avoid similar pitfalls in the creation of a new system based on the AOSpine classification system.

Conclusion

The AOSpine classification system would greatly benefit from the incorporation of a numerical scoring system to allow the treating clinician to more effectively evaluate all spinal injuries with respect to optimal treatment strategy. This may be particularly useful for injuries such as lumbar burst fractures, which treatment methods may be controversial. While the AOSpine classification system successfully integrates the Magerl and TLICS systems, the addition of a scoring system, similar to that seen in TLICS, would make the AOSpine classification system more practical and functional in regards to determining the most beneficial and effective treatment for the patient.

References

1. Böhler L (1956) Mechanisms of fracture and dislocation of the spine in the treatment of fractures. Fractures and dislocation of the spine 1: 300-329.

2. NICOLL EA (1949) Fractures of the dorso-lumbar spine. J Bone Joint Surg Br 31B: 376-394.

3. Holdsworth F (1963) Fractures, dislocations, and frature-dislocations of the spine. J Done Joint Surg Br 45: 6-20.

4. Holdsworth F (1970) Fractures, dislocations, and fracture-dislocations of the spine. J Bone Joint Surg Am 52: 1534-1551.

5. Denis F (1983) The three column spine and its significance in the classification of acute thoracolumbar spinal injuries. Spine (Phila Pa 1976) 8: 817-831.

6. Vaccaro AR, Lehman RA Jr, Hurlbert RJ, Anderson PA, Harris M, et al. (2005) A new classification of thoracolumbar injuries: the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. Spine (Phila Pa 1976) 30: 2325-2333.

7. McAfee PC, Yuan HA, Fredrickson BE, Lubicky JP (1983) The value of computed tomography in thoracolumbar fractures. An analysis of one hundred consecutive cases and a new classification. J Bone Joint Surg Am 65: 461-473.

8. Ferguson RL, Allen BL Jr (1984) A mechanistic classification of thoracolumbar spine fractures. Clin Orthop Relat Res : 77-88.

9. Chapman JR, Norvell DC (2009) Spine Classifications and Severity Measures. AOSpine, Dübendorf, Switzerland:.

10. Mirza SK, Mirza AJ, Chapman JR, Anderson PA (2002) Classifications of thoracic and lumbar fractures: rationale and supporting data. J Am Acad Orthop Surg 10: 364-377.

11. Wood KB, Khanna G, Vaccaro AR, Arnold PM, Harris MB, et al. (2005) Assessment of two thoracolumbar fracture classification systems as used by multiple surgeons. J Bone Joint Surg Am 87: 1423-1429.

12. Whang PG, Vaccaro AR, Poelstra KA, Patel AA, Anderson DG, et al. (2007) The influence of fracture mechanism and morphology on the reliability and validity of two novel thoracolumbar injury classification systems. Spine (Phila Pa 1976) 32: 791-795.

13. Vaccaro AR, Oner C, Kepler CK, Dvorak M, Schnake K, et al. (2013) AOSpine thoracolumbar spine injury classification system: fracture description, neurological status, and key modifiers. Spine (Phila Pa 1976) 38: 2028-2037.

14. Magerl F, Aebi M, Gertzbein SD, Harms J, Nazarian S (1994) A comprehensive classification of thoracic and lumbar injuries. Eur Spine J 3: 184-201.

15. Koh YD, Kim DJ, Koh YW (2010) Reliability and Validity of Thoracolumbar Injury Classification and Severity Score (TLICS). Asian Spine J 4: 109-117.

16. Lewkonia P, Paolucci EO, Thomas K (2012) Reliability of the thoracolumbar injury classification and severity score and comparison with the denis classification for injury to the thoracic and lumbar spine. Spine (Phila Pa 1976) 37: 2161-2167.

17. Joaquim AF, Fernandes YB, Cavalcante RA, Fragoso RM, Honorato DC, et al. (2011) Evaluation of the thoracolumbar injury classification system in thoracic and lumbar spinal trauma. Spine (Phila Pa 1976) 36: 33-36.

18. Patel AA, Whang PG, Brodke DS, Agarwal A, Hong J, et al. (2007) Evaluation of two novel thoracolumbar trauma classification systems. Indian J Orthop 41: 322-326.

19. Whang PG, Vaccaro AR, Poelstra KA, Patel AA, Anderson DG, et al. (2007) The influence of fracture mechanism and morphology on the reliability and validity of two novel thoracolumbar injury classification systems. Spine (Phila Pa 1976) 32: 791-795.

20. Moore TA1, Bransford RJ, France JC, Anderson J, Steinmetz MP, et al. (2014) Low lumbar fractures: does thoracolumbar injury classification and severity score work? Spine (Phila Pa 1976) 39: E1021-1025.

21. Mattei TA, Hanovnikian J, D HD (2014) Progressive kyphotic deformity in comminuted burst fractures treated non-operatively: the Achilles tendon of the Thoracolumbar Injury Classification and Severity Score (TLICS). European spine journal.

22. Joaquim AF, Ghizoni E, Tedeschi H, Batista UC, Patel AA (2014) Clinical results of patients with thoracolumbar spine trauma treated according to the Thoracolumbar Injury Classification and Severity Score. J neurosurg Spine 20: 562-567.

23. Reinhold M, Audigé L, Schnake KJ, Bellabarba C, Dai LY, et al. (2013) AO spine injury classification system: a revision proposal for the thoracic and lumbar spine. Eur Spine J 22: 2184-2201.

24. Rihn JA, Yang N, Fisher C, Saravanja D, Smith H, et al. Using magnetic resonance imaging to accurately assess injury to the posterior ligamentous complex of the spine: a prospective comparison of the surgeon and radiologist. J neurosurg Spine 12: 391-396.

25. Vaccaro AR, Rihn JA, Saravanja D, Anderson DG, Hilibrand ASD, et al. (2009) Injury of the posterior ligamentous complex of the thoracolumbar spine: a prospective evaluation of the diagnostic accuracy of magnetic resonance imaging. Spine 34: E841-847.