AOSpine thoracolumbar spine injury classification system: Fracture description, neurological status, and key modifiers

ARTICLE *in* SPINE · NOVEMBER 2013

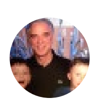
Impact Factor: 2.45 · DOI: 10.1097/BRS.0b013e3182a8a381

15 AUTHORS, INCLUDING:

Carlo Bellabarba
University of Washington Seattle

Frank Kandziora
Berufsgenossenschaftliche Unfallklinik Fr…

Jens R Chapman
University of Washington Seattle



Luiz Vialle
Pontifícia Universidade Católica do Paran…

Spine

SPINE Volume 38, Number 23, pp 2028-2037

DIAGNOSTICS

AOSpine Thoracolumbar Spine Injury Classification System

Fracture Description, Neurological Status, and Key Modifiers

Alexander R. Vaccaro, MD, PhD,* Cumhur Oner, MD, PhD,† Christopher K. Kepler, MD, MBA,* Marcel Dvorak, MD,‡ Klaus Schnake, MD,§ Carlo Bellabarba, MD,¶ Max Reinhold, MD,‖ Bizhan Aarabi, MD,** Frank Kandziora, MD, PhD,§ Jens Chapman, MD,†† Rajasekaran Shanmuganathan, MD, PhD,‡‡ Michael Fehlings, MD, PhD,§§ Luiz Vialle, MD, PhD,¶¶ and for the AOSpine Spinal Cord Injury & Trauma Knowledge Forum

Study Design. Reliability and agreement study, retrospective case series.

Objective. To develop a widely accepted, comprehensive yet simple classification system with clinically acceptable intra- and interobserver reliability for use in both clinical practice and research.

Summary of Background Data. Although the Magerl classification and thoracolumbar injury classification system (TLICS) are both well-known schemes to describe thoracolumbar (TL) fractures, no TL injury classification system has achieved universal international adoption. This lack of consensus limits communication between clinicians and researchers complicating the study of these injuries and the development of treatment algorithms.

Methods. A simple and reproducible classification system of TL injuries was developed using a structured international consensus process. This classification system consists of a morphologic classification of the fracture, a grading system for the neurological status, and description of relevant patient-specific modifiers. Forty cases with a broad range of injuries were classified independently twice by group members 1 month apart and analyzed for classification reliability using the Kappa coefficient (κ).

Results. The morphologic classification is based on 3 main injury patterns: type A (compression), type B (tension band disruption), and type C (displacement/translation) injuries. Reliability in the identification of a morphologic injury type was substantial (κ = 0.72).

Conclusion. The AOSpine TL injury classification system is clinically relevant according to the consensus agreement of our international team of spine trauma experts. Final evaluation data showed reasonable reliability and accuracy, but further clinical validation of the proposed system requires prospective observational data collection documenting use of the classification system, therapeutic decision making, and clinical follow-up evaluation by a large number of surgeons from different countries.

Key words: spinal injury classification, thoracolumbar, consensus development, agreement study, reliability, accuracy.

Level of Evidence: 4

Spine 2013;38:2028–2037

From the *Thomas Jefferson University and The Rothman Institute, Philadelphia, PA; †University Medical Center, Utrecht, the Netherlands; ‡Vancouver General Hospital, Vancouver, British Columbia, Canada; §Unfallklinik Frankfurt am Main, Frankfurt, Germany; ¶University of Washington, Seattle, WA; ‖Medical University Innsbruck, Innsbruck, Austria; **University of Maryland School of Medicine, Baltimore, MD; ††Harborview Medical Center, Seattle, WA; ‡‡Ganga Hospital, Coimbatore, Tamil Nadu, India; §§University of Toronto, Toronto, Ontario, Canada; and ¶¶Catholic University, Curitiba, Brazil.

Acknowledgment date: April 3, 2013. Revision date: June 14, 2013. Acceptance date: June 24, 2013.

The manuscript submitted does not contain information about medical device(s)/drug(s).

The AOSpine funds were received in support of this work. AOSpine is a clinical division of the AO Foundation—an independent medically guided not-for-profit organization. The AO has a strong financial independence thanks to the foundations endowment. The annual operating activities are financed through 3 pillars: Collaboration and support agreements with DePuy Synthes and other industrial partners, return on own financial assets, and other third party income (e.g., participant fees, R&D projects, memberships). The AOSpine Knowledge Forums are pathology-focused working groups acting on behalf of AOSpine in their domain of scientific expertise. Each forum consists of a steering committee of up to 10 international spine experts who meet biannually to discuss research, assess the best evidence for current practices, and formulate clinical trials to advance their field of spine expertise. Authors are compensated for their travel and accommodation costs. Study support is provided directly through AOSpine's Research department and AO's Clinical Investigation and Documentation unit. There are no other institutional subsidies, corporate affiliations, or funding sources supporting this work unless clearly documented and disclosed.

Relevant financial activities outside the submitted work: board membership, consultancy, grants, payment for lecture, payment for manuscript preparation, patents, royalties, stocks.

Address correspondence and reprint requests to Alexander R. Vaccaro, MD, PhD, 925 Chestnut St, Fifth Flr, Philadelphia, PA 19107; E-mail: alexvaccaro3@aol.com

DOI: 10.1097/BRS.0b013e3182a8a381

Classification of spinal fractures to facilitate communication and encourage optimal treatment protocols has long been a focus of the spine community. Many classification systems have been proposed but none has achieved universal adoption. Proposed systems have used diverse injury characteristics as the basis for classification such as inferred mechanism of injury,[1] bony morphology,[2–5] anatomic determinants of fracture stability,[5,6] and neurological status.[3,6] In particular, the contributions of McAfee *et al*. lead to an understanding of the association between fracture morphology and stability, providing much of the intellectual basis for later schemes which used morphology to describe fracture stability and treatment recommendations.[7–9] Of the morphology-based classification systems, the Comprehensive Classification scheme proposed by Magerl *et al*[4] is arguably the most systematic and detailed. The Magerl classification includes a comprehensive description of fracture anatomy and was intended to follow a hierarchical system in which successive grades represent increasing fracture severity, instability, and consequently an inferred increased risk of neurological injury, by comprehensively describing subdivisions of injury variants. Criticized for being overly complex, the Magerl system did not give formal consideration to the neurological injury or other clinical factors which may guide surgeon decision making,[6,10] concepts increasingly embraced as classification systems are expected to provide prognostic, and treatment guidance. Furthermore, the Magerl classification has neither been clinically validated nor revised to improve its reliability and clinical applicability.[11,12]

In contrast to the Magerl system, the thoracolumbar injury classification system (TLICS) evaluates the neurological status, integrity of the posterior ligamentous complex (PLC), and injury morphology of each patient using descriptive categories.[6] TLICS also aims to guide treatment decision using a scoring system, which assigns point values based on neurological status, integrity of the PLC, and morphology. Point totals are then used to recommend surgical or nonsurgical treatment, or the point total is indeterminate, and the treating surgeons must use their clinical judgment. Although inclusion of neurological status in the scheme may increase the clinical relevance of this system, the TLICS has also met with several criticisms. The reproducibility and feasibility of evaluating PLC integrity using magnetic resonance imaging (MRI) has been questioned.[13,14] Also, the chosen severity scoring system guiding treatment may be a culture- or region-specific decision and may not reflect global surgical preferences or the most rational approach to treatment.

The AOSpine Trauma Knowledge Forum, an international group of academic spine surgeons, was tasked to develop and validate a classification system incorporating both fracture morphology and clinical factors relevant for surgical decision making, such as the presence of neurological deficits. The goal of this effort was to develop a widely accepted, comprehensive yet simple classification system with clinically acceptable intra- and interobserver reliability to be used for clinical practice and research purposes.

MATERIALS AND METHODS

The methodological background of the entire process and the 4 spine regions (upper cervical, lower cervical, thoracolumbar [TL], and sacral) to be described by this classification system have been separately described in detail along with an earlier iteration of the classification system.[15] A workgroup of the AOSpine (AOSpine Classification Group) has systematically assessed and revised the Magerl classification using an AOSpine database of more than 750 spinal trauma cases with digital imaging and communications in medicine images to develop a rational, simple, and reproducible morphologic classification. This workgroup conducted evaluation sessions to assess the reliability and accuracy and identified areas of disagreement, which required further refinement until a unanimous consensus was reached regarding classification details and application of the system, and adequate reliability was achieved.

Seven face-to-face meetings and 5 evaluation sessions were ultimately necessary to achieve a consensus system for grading TL fracture morphology. The results of the final session with respect to reliability analysis of fracture morphology are presented; during the final session the concepts of neurological and patient-specific modifiers were incorporated with the eventual goal of predicting treatment approaches and prognosis in addition to describing morphology. The case series during this last final session included 40 cases representing a random selection of TL injuries from one author's practice across all grades of injury and neurological status. A consecutive series could not be used for this study as the less severe grades of injuries predominate on the basis of an incidence, and the resulting reliability analysis would not accurately describe application to a complete range of injury morphology. In total, 9 fellowship-trained spine surgeons with experience in spinal trauma graded the cases. A second round of grading 1 month after the first round was performed after the case order had been scrambled using a random number generator.

Cases were graded by injury type (A, B, C). Because each patient could potentially have more than one injured spinal level, the level of injury to be graded was designated when imaging demonstrated multiple injuries. For type A injuries, only cases with single vertebral body injury (disregarding the B and C coding) were included to ensure that the surgeons were assessing the same injury. Verbal descriptions of injury patterns were combined with standardized iconic images of each injury type allowing both a rigorous visual and linguistic descriptive understanding of each injury pattern. Verbal descriptors were not incorporated into the reliability analysis.

For type B and type C injuries, concurrent type A or type B injuries at the same level were graded by readers, but only the most severe injury was considered for the purposes of data analysis.

Statistical analysis used the Kappa coefficient (κ) to assess the reliability of the classification system among different observers (interobserver agreement) and the reproducibility for the same observer on separate occasions (intraobserver reproducibility). The coefficients were interpreted using the Landis and Koch grading system,[16] which defines κ of less than 0.2 as slight agreement or reproducibility, between 0.2 and 0.4 as fair agreement or reproducibility, between 0.4 and 0.6 as moderate agreement or reproducibility, between 0.6 and 0.8 as substantial reliability or reproducibility, and more than 0.8 as excellent reliability or reproducibility. κ coefficients were calculated for most severe injury type (*i.e.*, A, B, or C), subtype (*e.g.*, A0, A1, A2, A3, or A4), and neurological status by history. Fractures categorized by at least one assessor as a type A fracture or as a type B fracture were included in a subgroup analysis for intrarater reproducibility of subtypes for type A and type B injuries.

RESULTS

This classification is based on the evaluation of 3 basic parameters:

1. Morphologic classification of the fracture
2. Neurological status
3. Clinical modifiers

Morphologic Classification

Similar to the Magerl system,[4] successive injury types indicate ascending severity of injury: Three basic types are identified on the basis of the mode of failure of the spinal column (Figure 1), and the algorithm to identify fractures is described in Supplemental Digital Content Appendix 1, available at http://links.lww.com/BRS/A810.

- *Type A:* Compression injuries.
- *Type B:* Failure of the posterior or anterior tension band without evidence of either gross translation or the potential for gross translation.
- *Type C:* Failure of all elements leading to dislocation or displacement in any plane or complete disruption of a soft-tissue hinge even in the absence of translation.

GRADING OF INJURIES

Type A injuries may affect a single vertebral body in isolation or occur in combination with type B or type C injuries. B2, B3, and type C injuries affect a motion segment and are coded accordingly by motion segment (*e.g.*, T12–L1), whereas A and B1 injuries are coded by the single vertebral level they affect (*e.g.*, L2). The colloquial term for the injury (*e.g.*, burst, compression fracture, and distraction extension injury) may be listed after the alphanumeric designation to increase acceptance and usage among those surgeons who are more familiar with these descriptors, although it is recognized this is somewhat redundant (see Figure 2, Supplemental Digital Content available at http://links.lww.com/BRS/A812). Multilevel injuries should be classified separately and listed according to declining severity. When injuries of the same subtype are present, the injuries will be described in order of cranial to caudal location. This system obligates treating surgeons to examine all TL motion segments carefully to accurately and comprehensively describe the injury accurately and comprehensively. Clinical cases are included in Supplemental Digital Content Appendix 2 available at http://links.lww.com/BRS/A811.

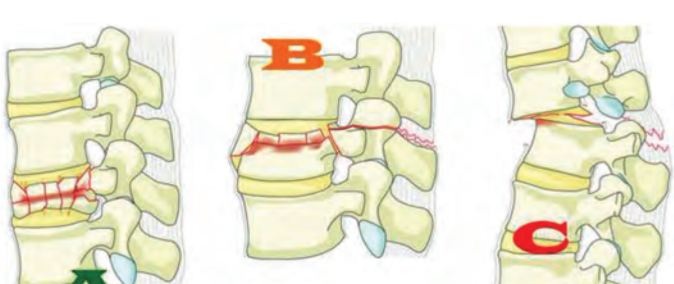


Figure 1. *The 3 basic types*—Type A: Compression injuries. Failure of anterior structures under compression with intact tension band. Type B: Failure of the posterior or anterior tension band. Type C: Failure of all elements leading to dislocation or displacement.

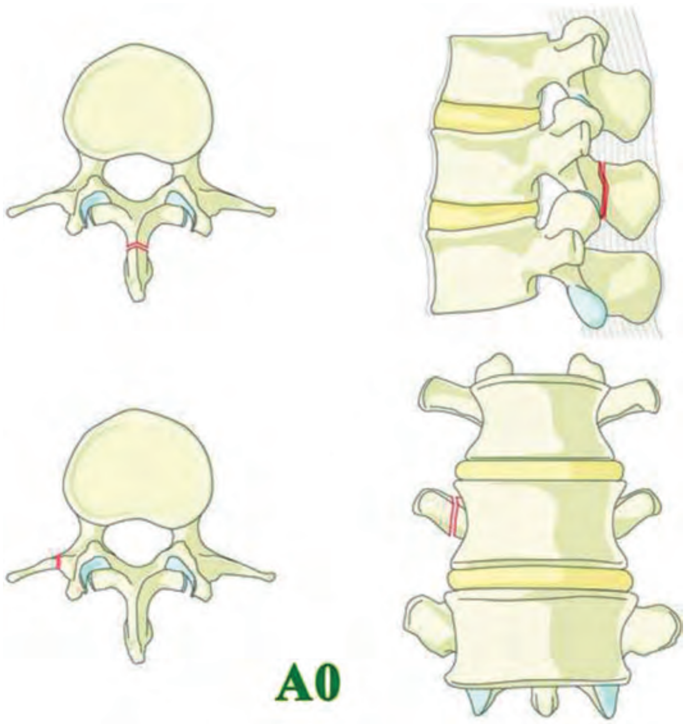


Figure 3. *Subtype A0—Minor injuries*: Injuries such as transverse process or spinous process fractures, which do not compromise the mechanical integrity of the spinal column. Figure 3 demonstrates schematic drawing of this injury while Figure 3.2 available at Supplemental Digital Content http://links.lww.com/BRS/A813 shows a CT scan of a patient with this injury. CT indicates computed tomography.

Type A Injuries: Compression Injuries of the Vertebral Body

Type A injuries involve the anterior elements (vertebral body and/or disc), and this type includes clinically insignificant injuries to the elements such as transverse or spinous process fractures. More severe type A injuries involve vertebral body burst fractures with retropulsion of the posterior vertebral body without disruption of the PLC and without any translation/displacement. Type A injuries are further divided into 5 subtypes.

- *Subtype A0* either designates no fracture of the vertebra or clinically insignificant fractures of the spinous or transverse processes as shown in Figure 3; Figure 3.2, Supplemental Digital Content available at http://links.lww.com/BRS/A813. *Comment:* Whether this designates lack of any visualized injury or a clinically insignificant injury, there is no concern for mechanical instability or a neurological deficit.

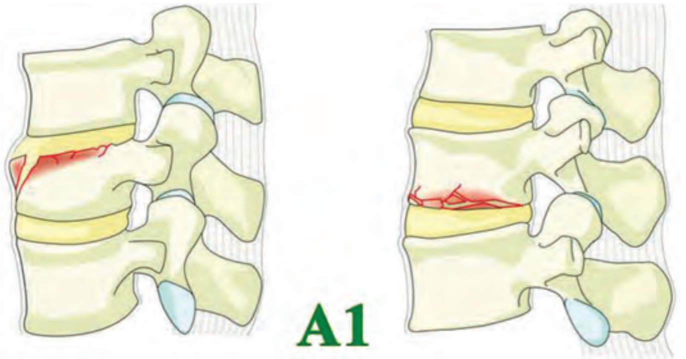


Figure 4. *Subtype A1—Wedge Compression*: Fracture of a single endplate without involvement of the posterior wall of the vertebral body. Vertebral canal is intact. Figure 4 demonstrates schematic drawing of this injury while Figure 4.2 available at Supplemental Digital Content http://links.lww.com/BRS/A814 shows a CT scan of a patient with this injury. CT indicates computed tomography.

- *Subtype* A1 injuries are wedge compression or impaction fractures with fracture of a single endplate without involvement of the posterior wall of the vertebral body as demonstrated in Figure 4; Figure 4.2, Supplemental Digital Content available at http://links.lww.com/BRS/A814.
- *Subtype* A2 injuries are split- or pincer-type fractures in which the fracture line involves both endplates but does not involve the posterior vertebral wall as shown in Figure 5; Figure 5.2, Supplemental Digital Content available at http://links.lww.com/BRS/A815.
- *Subtype A3* injuries are vertebral fractures affecting a single endplate with any involvement of the posterior vertebral wall and the spinal canal as shown in Figure 6; Figure 6.2, Supplemental Digital Content available at http://links.lww.com/BRS/A816. The compressive forces may also result in increased interpedicular distances and vertical (greenstick-like) fractures of the lamina. The integrity of the posterior tension band is maintained and there is no vertebral translation. Injuries with ligamentous disruption of the posterior tension band should be primarily classified as B2 injuries. A3 fractures of the body that involve an axial-plane horizontal fracture through the posterior elements (as opposed to the vertical fracture described in earlier text) disrupts the stability of the spine, and such an injury should be classified as a type B injury.
- *Subtype A4* injuries, shown in Figure 7; Figure 7.2, Supplemental Digital Content available at http://links.lww.com/BRS/A817, are vertebral body fractures involving both endplates as well as the posterior wall. Similar to A3 injuries, these may be associated with vertical fracture lines of the lamina but without disruption of the posterior tension band. Injuries with ligamentous disruption of the posterior tension band should be primarily classified as B2 injuries. A4 injuries are similar to A3 injuries but involve both endplates. Split fractures that also involve the posterior vertebral body are included in this group. A4 fractures of the body that involve an axial-plane horizontal fracture through the posterior elements (as opposed to the vertical fracture described in earlier text) disrupts the stability of the spine, and such an injury should be classified as a type B injury.

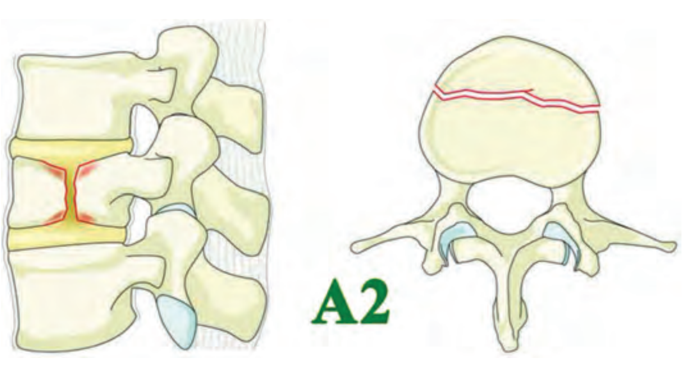


Figure 5. *Subtype A2—Split or pincer-type*: Fracture of both endplates without involvement of the posterior wall of the vertebral body. Figure 5 demonstrates schematic drawing of this injury while Figure 5.2 available at Supplemental Digital Content http://links.lww.com/BRS/A815 shows a CT scan of a patient with this injury. CT indicates computed tomography.

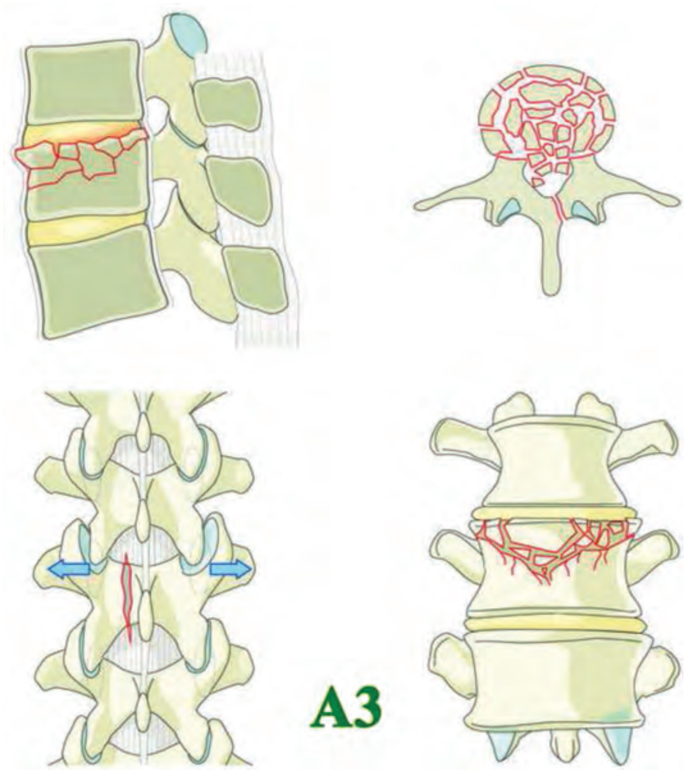


Figure 6. *Subtype A3—Incomplete burst:* Fracture with any involvement of the posterior wall of the vertebral body. Only a single endplate fractured. Vertical fracture of the lamina is usually present and does not indicate a tension band failure. Figure 6 demonstrates schematic drawing of this injury while Figure 6.2 available at Supplemental Digital Content http://links.lww.com/BRS/A816 shows a CT scan of a patient with this injury. CT indicates computed tomography.

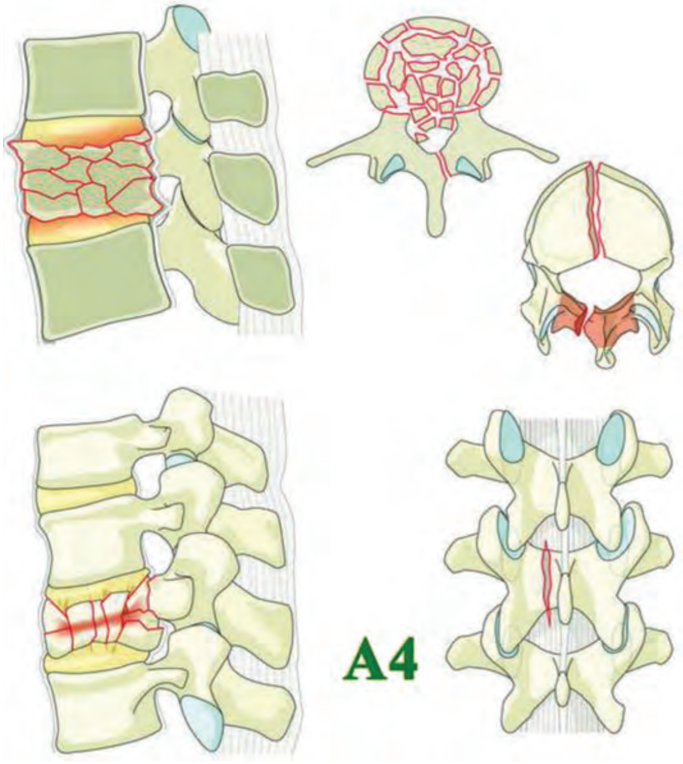


Figure 7. *Subtype A4—Complete burst:* Fracture with any involvement of the posterior wall of the vertebral body and both endplates. Vertical fracture of the lamina is usually present and does not indicate a tension band failure. Figure 7 demonstrates schematic drawing of this injury while Figure 7.2 available at Supplemental Digital Content http://links.lww.com/BRS/A817 shows a CT scan of a patient with this injury. CT indicates computed tomography.

Type B Injuries: Tension Band Injury

Type B injuries affect either anterior or posterior tension band. These injuries may be seen in combination with type A fractures of the vertebral body. They are further divided in 3 subgroups.

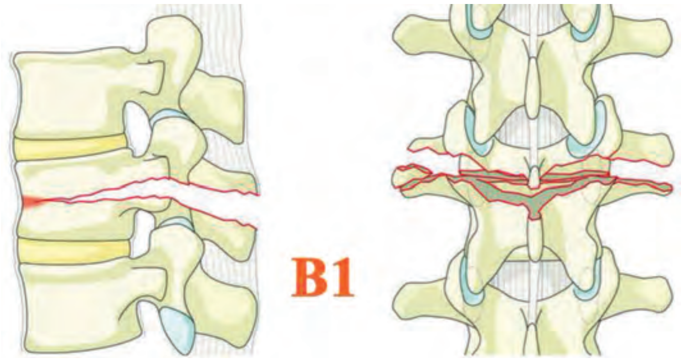


Figure 8. *Subtype B1—Monosegmental bony posterior tension band injury:* Transosseous failure of the posterior tension band. The classical "chance fracture." Figure 8 demonstrates schematic drawing of this injury while Figure 8.2 available at Supplemental Digital Content http://links.lww.com/BRS/A818 shows a CT scan of a patient with this injury. CT indicates computed tomography.

- *Subtype B1* injuries are monosegmental osseous failure of the posterior tension band extending into the vertebral body, known as "chance" fractures, as shown in Figure 8; Figure 8.2, Supplemental Digital Content available at http://links.lww.com/BRS/A818. Unlike the B2 subtype that always affects an intervertebral level, the B1 subtype affects a single vertebral body level. The fracture may extend through the pedicle and exit from the posterior aspect of the pars interarticularis into the posterior soft tissues or extend through the pedicle through the spinous process before exiting into the soft tissue posteriorly.
- *Subtype B2* injuries demonstrate a disruption of the posterior tension band with or without osseous involvement, shown in Figure 9; Figure 9.2, Supplemental Digital Content available at http://links.lww.com/BRS/A819. Any associated vertebral body compression fracture should be specified separately according to the corresponding type A subdivision. In particular, patients with burst fractures demonstrated to have disruption of the PLC on MR image should be described as having a B2 injury with either an A3 (incomplete burst) or A4 (complete burst) vertebral body injury.
- *Subtype B3* injuries disrupt the anterior longitudinal ligament (ALL) that serves as the anterior tension band of the spine, preventing hyperextension. The injury may pass through either the intervertebral disc or through the vertebral body

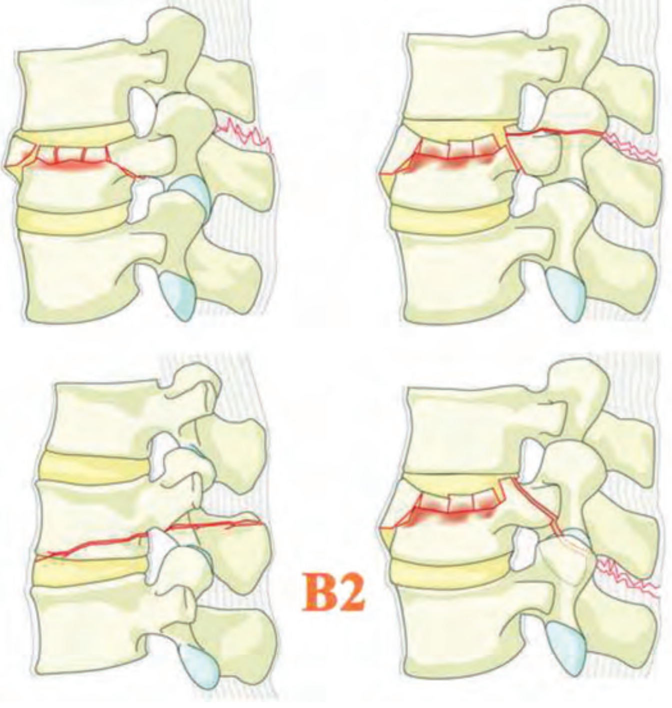


Figure 9. *Subtype B2—Posterior tension band disruption:* Bony and/or ligamentary failure of the posterior tension band together with a type A fracture. type A fracture should be classified separately. This example should be classified as: T12–L1 "type B2" with T12 "A4" according to the combination rules. Figure 9 demonstrates schematic drawing of this injury while Figure 9.2 available at Supplemental Digital Content http://links.lww.com/BRS/A819 shows a CT scan of a patient with this injury. CT indicates computed tomography.

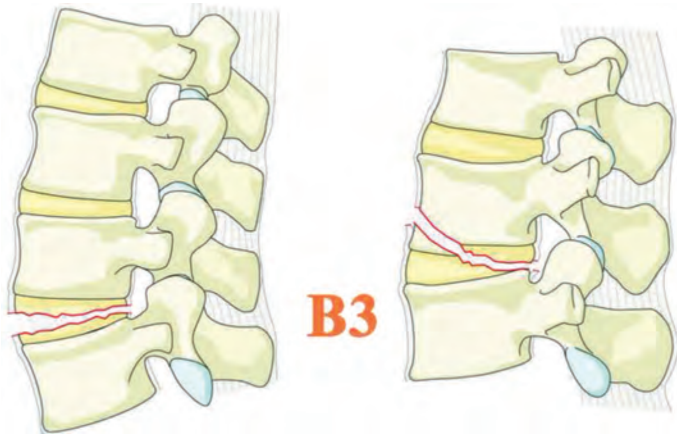


Figure 10. *Subtype B3—Hyperextension injury:* Injury through the disc or vertebral body leading to a hyperextended position of the spinal column, which is commonly seen in ankylotic disorders. Anterior tension band, notably the ALL is ruptured but there is a posterior hinge preventing further displacement. Figure 10 demonstrates schematic drawing of this injury while Figure 10.2 available at Supplemental Digital Content http://links.lww.com/BRS/A820 shows a CT scan of a patient with this injury. CT indicates computed tomography. CT indicates computed tomography.

itself (particularly in the ankylosed spine), but there is an intact posterior element hinge preventing gross displacement. Postinjury imaging often demonstrates a hyperextended malalignment as seen in Figure 10; Figure 10.2, Supplemental Digital Content available at http://links.lww.com/BRS/A820. Clear and complete disruption of the posterior hinge removes the barrier to translation, and then, the injury should be considered a type C injury with a B descriptor, even in the absence of displacement/translation at the time of injury.

Type C Injuries: Displacement/Translational Injury

Type C injuries are characterized by displacement beyond physiological range of the cranial and caudal parts of the spinal column in any plane demonstrated in Figure 11; Figure 11.2, Supplemental Digital Content available at http://links.lww.com/BRS/A821. Type C injuries also occur in the presence of distraction of both the anterior and posterior vertebral elements without any remaining intact anterior or posterior structure, there may be complete separation of the vertebral elements. Any associated vertebral body fracture should be specified separately (*e.g.*, A0, A1, A2, A3, A4). Any associated tension band injuries should be specified separately (*e.g.*, B1, B2, B3), if possible to provide greater insight into injury morphology.

GRADING OF NEUROLOGICAL DEFICITS

Neurological status is graded according to a 5-part system:

- *N0* is used to designate patients who are neurologically intact.
- *N1* means that a patient had a transient neurological deficit, which is no longer present.
- *N2* denotes patients with symptoms or signs of radiculopathy.

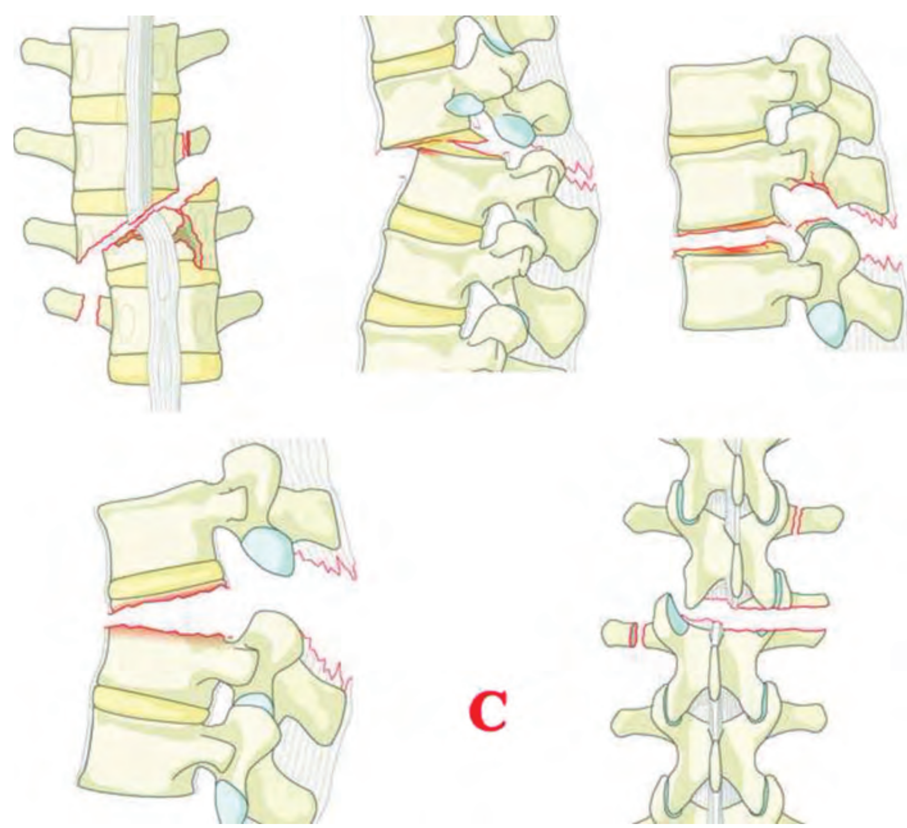


Figure 11. *Type C—Translation/displacement:* There are no subtypes as because of the dissociation between cranial and caudal segments various configurations are possible in different images, which are not relevant. Is combined with subtypes of A to denote the associated vertebral body fractures if necessary. Figure 11 demonstrates schematic drawing of this injury while Figure 11.2 available at Supplemental Digital Content http://links.lww.com/BRS/A821 shows a CT scan of a patient with this injury. CT indicates computed tomography.

- *N3* incomplete spinal cord injury or cauda equina injury.
- *N4* complete spinal cord injury (American Spinal Injury Association grade A[17]).

NX is used to designate patients who cannot be examined because of head injury or another condition, which limits their ability to complete a neurological examination such as intoxication, multiple trauma, or intubation/sedation.

Case-Specific Modifiers

Two additional modifiers were thought to be important enough for inclusion but would not be relevant to every case and used on an as-needed basis to assist the physician in deciding treatment.

M1 is used to designate fractures with an indeterminate injury to the tension band based on spinal imaging such as MRI or clinical examination. This modifier is important for identifying those injuries that seem stable from a bony standpoint for which ligamentous insufficiency may help determine whether operative stabilization is a consideration.

M2 is used to designate a patient-specific comorbidity, which might argue either for or against surgery for those patients with relative indications for surgery. Examples of a M2 modifier include such disorders but not limited to ankylosing spondylitis, rhematologic conditions, diffuse idiopathic skeletal hyperostosis, osteopenis/porosis, or burns affecting the skin overlying the injured spine.

Spine Injury Score

A Spine Injury Score is an integral part of the TLICS system. A suggested scoring system needs to be reflective of the natural healing process of a spinal injury but also sensitive to the experience and expertise of the region or culture in which it is used, reflecting a society's value on cost as well as the timeliness of function and rehabilitation. This classification system will be mated with a severity scoring system that will be validated throughout the international spinal community and may be different for different regions depending on the preferences of societal values.

Final Evaluation Session

Interobserver Reliability

On the basis of the surgeons' classification of fracture morphology, the proportion of injuries in the random sample series demonstrating a type A fracture of a single vertebra was 54%. The percentage of cases reported on average to have a type B injury was 24%, whereas the type C injuries represented 22% of cases on average.

Full agreement among all surgeons was achieved when classifying the type of injury for 14 of 40 (35%) of the cases; the overall κ coefficient for all cases was 0.64. When comparing grading by fracture type regardless of subtype (A/B/C), investigators classified fractures unanimously in 24 of 40 cases (60%). These cases included 16 type A fractures, 3 type B fractures, and 5 type C fractures. The κ statistic for overall agreement on grading by fracture type without regard to subtype was 0.72.

TABLE 1. Distribution of Injuries and κ Coefficients of Reliability for Each TL Injury Type

Type	n* (%)	κ
A Compression fractures		
0 No injury/process fracture	44 (5)	1.0
1 Wedge/impaction	95 (11)	0.59
2 Split/pincer type	61 (7)	0.50
3 Incomplete burst	107 (12)	0.45
4 Complete burst	164 (19)	0.58
Overall type A		0.72
B Tension band injuries		
1 Posterior transosseous disruption	70 (8)	0.65
2 Posterior ligamentous disruption	98 (11)	0.34
3 Anterior ligamentous disruption	48 (5)	0.41
Overall type B		0.58
C Translation injuries		
	193 (22)	
Overall type C		0.70

The sample of type A fractures included all cases with a single compression fracture and with or without type B or type C injury within the random sample of 110 cases. The sample of type B and type C injuries included all such injuries identified within the complete TL case series. This table includes results from 9 raters.

**Estimation of case distribution by the grading surgeons.*

κ indicates Kappa.

κ values describing interobserver agreement were 0.72 for type A injuries, 0.58 for type B injuries, and 0.7 for type C injuries. The lowest level of agreement for specific subtypes was for fracture type B2 (κ = 0.34) and B3 (κ = 0.41) (Table 1).

Intraobserver Reliability

All raters had substantial to excellent reproducibility results for TL morphology classification with an average κ value of 0.77 (range, 0.6–0.97). Reproducibility of fracture type without regard for subtype was excellent with κ = 0.85 (range, 0.75–0.96). Intrarater reproducibility for subtypes of type A and B fractures demonstrated κ = 0.72 and κ = 0.43.

DISCUSSION

In this article, we describe the development of a TL spinal fracture classification system that accounts simply for the various patterns of spinal fracture and soft-tissue injury, the extent of neurological deficit and the presence or absence of key medical comorbidities. A multitude of spine classification and severity measures have been developed but none has led to a universal TL spinal injury classification system that is widely accepted and used. No single classification has been able to simultaneously describe injury severity and pathomorphology while considering all clinical, neurological, and radiological characteristics[18] relevant to clinical decision making.

There is little guidance for the development of spinal injury classifications in general, and most past efforts have heavily used expert opinion.[19] We think strongly that morphological characteristics that can be reliably and reproducibly identified must be the backbone of any fracture classification system. For this reason, the presented classification provides clear injury characteristics of the vertebral column, and the system was designed to be primarily based around features identifiable using CT scan, a widely available imaging modality at most trauma centers.[20] The proposed classification system presents distinct and specific morphologic injury characteristics that can be used to distinguish one injury subgroup clearly from another primarily using CT scan. Importantly, the proposed classification scheme goes beyond fracture morphology to acknowledge the relevance of patient comorbidities and neurological status in making treatment decisions, reflecting the contributions of TLICS.[6]

MRI can be useful for diagnosing subtle PLC injury, particularly in those situations where fracture displacement on presentation is not representative of maximal displacement at the time of injury. Furthermore, MRI is often helpful in determining the location and severity of neurological compromise and identifying injury to nonbony structures; MRI shows higher sensitivity, specificity, and accuracy in distinguishing ligamentous lesions *versus* CT[21,22] and may reduce the risk of failure to diagnose a PLC injury and associated late deformity.[23,24] We recognize the limitations of MRI, namely the relatively poor reliability associated with identification of PLC injury, and we acknowledge that a classification system heavily dependent on MRI would be unlikely to gain widespread usage in the developing world. Prospective study across a variety of hospitals with differing access to advanced imaging has been initiated to demonstrate whether there is a discernible difference in treatment pattern or outcome based on the availability of MRI.

The importance of imaging in a spinal injury classification system cannot be overstated. Improvements in image quality and the development of multiplanar imaging have greatly improved our understanding of fracture morphology[20]; this increased anatomic precision should be reflected in any classification system if it is to remain clinically relevant and improve upon the characterization of morphology proposed by early classification systems that relied entirely on plain radiographs.[25] In the vast majority of cases, accurate classification is possible with CT scan and/or plain radiographs. In the current scheme, MRI may be used to demonstrate disruption of the anterior or posterior tension band, demonstrating that an injury is at least a type B or may be used to demonstrate that the posterior hinge is disrupted, and that an extension-distraction injury is actually a type C injury. The M1 modifier is designed to give greater consideration for surgical intervention when the integrity of the PLC is indeterminate.

Similar to the manner in which PLC evaluation was incorporated into the classification scheme, other considerations, which potentially affect surgical decision making were also included in a qualitative manner without direct impact on the spine injury severity score. Rather than directly influence the need for surgery, we anticipate that the M2 patient-specific modifier will require case-by-case evaluation; further study will determine the relative value of these modifiers to the classification. In contrast, the patient's neurological status is critical for a complete assessment of the patient's functional status, and eventual prognosis and has been identified as one of the most important factors when making decisions about the need for surgery.[26] Given its critical importance, the neurological modifier is central to this classification proposal, along with fracture morphology, to guide the need for operative intervention.

Previously proposed classification schemes have disappointingly not reached the ideal mix of simplicity and comprehensiveness, a difficult balance that is necessary to achieve widespread adoption and application. Wood *et al*[10] studied the Denis[3] and the Magerl systems,[4] finding only moderate reliability and repeatability. Most concerning, repeated application of the Magerl system demonstrated that spine surgeons graded the same fractures differently 3 months after initial assessment 21% of the time.[10] This relatively low reliability may be partially attributable to the complexity of the Magerl classification scheme, which requires a high degree of familiarity with the system for correct application. Other studies of the interobserver reliability of the Magerl scheme reported κ coefficients of 0.33[23] and 0.62[11] for identification of the main injury types, below the corresponding value of 0.72 reported in this study.

Similarly, TLICS has been assessed for reliability of identification of fracture morphology, PLC injury, and treatment recommendation.[13,27,28] Two components of the TLICS scoring system, fracture morphology, and integrity of the PLC, were evaluated for interobserver reliability. Of these 2 factors, identification of PLC injury demonstrated lower interobserver reliability, achieving a κ value of 0.455. This difficulty in reliably identifying PLC injury was also demonstrated by Rihn *et al*[14] who reported a κ value of 0.58 by spine surgeons and 0.37 for musculoskeletal radiologists in identification of injury to the PLC using MRI. Rihn additionally reported that specificity of MR for identification of PLC injury was as low as 52% for some observers. Whang *et al*[13] found that identification of injury morphology had a substantially higher κ value of 0.626, whereas overall management decision demonstrated a κ value of 0.652. We hesitate to place undue importance on the agreement of management decision as this aspect of the analysis may be disproportionately influenced by close agreement regarding neurological status and its importance as a determinant of treatment, whereas the reproducibility of the other criteria is lower. Furthermore, there is disagreement in the international spine surgery community about guidelines for surgical intervention proposed by TLICS, rendering evaluation of the reliability of management decisions less clinically relevant. Nevertheless, until the introduction of the present system, TLICS remained the sole TL classification scheme that considered neurological status and advanced imaging and was influential in the design of the current system.

Evaluation of the morphological portion of the classification system demonstrated an interobserver κ coefficient of 0.64 for all fracture types and subtypes. The classification of both long bone and spinal fractures is often characterized by low reliability[11,23,29]; we consider this degree of agreement acceptable at this early stage in the development of this classification as it supersedes κ values published in relationship to previous systems to classify fracture morphology.[10,13] Careful analysis of cases where surgeons did not agree provides insight into problematic injuries. The unambiguous lack of involvement of the posterior structures clearly differentiates type A injuries from types B and C; this is borne out by the reliability data that suggest type A injuries are the most reliably identified type ($\kappa = 0.72$). Although almost all patients with type A injuries will be similarly graded by spine surgeons, a higher proportion of patients with either type B or type C injuries will be misdiagnosed. B1 injuries and B2 injuries are distinguished by the presence of a ligamentous component in the latter, a distinction that not may be easy to make in some cases. Another problematic injury type we identified involved an extension injury in an ankylosed spine. Imaging showed no displacement but some observers graded the injury as B3 and others as C. Despite the lack of displacement, the doubtful posterior hinge in a stiff spine makes this a highly unstable injury. Our definitions for type B and C injuries emphasizes the importance of the hinge for type B injuries and the lack of a hinge as a criteria for type C injuries, which have a high potential for further displacement regardless of the degree of displacement at the time of imaging. Concern over this distinction may be academic, however because we would expect that most patients with either extension-distraction injuries, and an ankylosed spine or translational injuries would be treated with surgical stabilization. Misdiagnosis of an injury as a less severe injury is of particular concern as such patients may receive treatment that does not sufficiently stabilize their spine.

CONCLUSION

We think that our proposed classification system represents a carefully developed, simple but comprehensive scheme, which simultaneously considers the inherent variability of spinal column injuries, all major modes of failure and clinical features such as neurological status that are critical to determining the need for surgery. To use this system in a prognostic manner that informs surgical decision making, broad and cross-cultural international prospective validation studies are needed, and investigations are already underway. Until this process is completed, no absolute recommendations about when surgery is mandatory can be provided, as this may be a reflection not only of a precise understanding of fracture stability, but also of cultural acceptance of surgical intervention reflecting the importance placed on immediate surgical stability and accelerated rehabilitation. Additional future goals of research include defining differences in outcome by fracture subgroup and delineating treatment algorithms that are fracture subgroup-specific when necessary.

➢ Key Points

- ❑ A new TLICS was developed by an international team that incorporates features of both the Magerl and TLICS.
- ❑ In addition to morphological description, this system considers neurological status and patient-specific modifiers that are important for surgical decision making.
- ❑ Grading of 40 cases by a group of spine surgeons experienced in spine trauma demonstrated substantial inter- and intraobserver reliability.
- ❑ Further study is necessary to validate this system on a broad international basis and to formulate a point system that will ultimately be used in treatment decisions.

Supplemental digital content is available for this article. Direct URL citations appearing in the printed text are provided in the HTML and PDF version of this article on the journal's web site (www.spinejournal.com).

References

1. Holdsworth F. Fractures, dislocations, and fracture-dislocations of the spine. *J Bone Joint Surg Am* 1970;52:1534–51.
2. Whitesides TE Jr. Traumatic kyphosis of the thoracolumbar spine. *Clin Orthop Relat Res* 1977;128:78–92.
3. Denis F. The three column spine and its significance in the classification of acute thoracolumbar spinal injuries. *Spine (Phila Pa 1976)* 1983;8:817–31.
4. Magerl F, Aebi M, Gertzbein SD, et al. A comprehensive classification of thoracic and lumbar injuries. *Eur Spine J* 1994;3:184–201.
5. McCormack T, Karaikovic E, Gaines RW. The load sharing classification of spine fractures. *Spine (Phila Pa 1976)* 1994;19:1741–4.
6. Vaccaro AR, Lehman RA, Jr Hurlbert RJ, et al. 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)* 2005;30:2325–33.
7. McAfee PC, Yuan HA, Fredrickson BE, et al. The value of computed tomography in thoracolumbar fractures: an analysis of one hundred consecutive cases and a new classification. *J Bone Joint Surg Am* 1983;65:461–73.
8. McAfee PC, Bohlman HH, Yuan HA. Anterior decompression of traumatic thoracolumbar fractures with incomplete neurological deficit using a retroperitoneal approach. *J Bone Joint Surg Am* 1985;67:89–104.
9. McAfee PC, Levine AM, Anderson PA. Surgical management of thoracolumbar fractures. *Instr Course Lect* 1995;44:47–55.
10. Wood KB, Khanna G, Vaccaro AR, et al. Assessment of two thoracolumbar fracture classification systems as used by multiple surgeons. *J Bone Joint Surg Am* 2005;87:1423–9.
11. Oner FC, Ramos LM, Simmermacher RK, et al. Classification of thoracic and lumbar spine fractures: Problems of reproducibility. A study of 53 patients using CT and MRI. *Eur Spine J* 2002;11:235–45.
12. Aebi M. Classification of thoracolumbar fractures and dislocations. *Eur Spine J* 2010;19(suppl 1):S2–7.
13. Whang PG, Vaccaro AR, Poelstra KA, et al. The influence of fracture mechanism and morphology on the reliability and validity of two novel thoracolumbar injury classification systems. *Spine (Phila Pa 1976)* 2007;32:791–5.
14. Rihn JA, Yang N, Fisher C, 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* 2010;12:391–6.

15. Reinhold M, Audigé L, Schnake KJ, et al. AO spine injury classification system: a revision proposal for the thoracic and lumbar spine [published online ahead of print March 19, 2013]. *Eur Spine J* doi:10.1007/s00586-013-2738-0.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
17. Maynard FM, Jr Bracken MB, Creasey G, et al. International standards for neurological and functional classification of spinal cord injury. American spinal injury association. *Spinal Cord* 1997;35:266–74.
18. Mirza SK, Mirza AJ, Chapman JR, et al. Classifications of thoracic and lumbar fractures: rationale and supporting data. *J Am Acad Orthop Surg* 2002;10:364–77.
19. van Middendorp JJ, Audige L, Hanson B, et al. What should an ideal spinal injury classification system consist of? A methodological review and conceptual proposal for future classifications. *Eur Spine J* 2010;19:1238–49.
20. Diaz JJ, Jr Cullinane DC, Altman DT, et al. Practice management guidelines for the screening of thoracolumbar spine fracture. *J Trauma* 2007;63:709–18.
21. Grunhagen J, Egbers HJ, Heller M, et al. Comparison of spine injuries by means of CT and MRI according to the classification of Magerl. *RöFo* 2005;177:828–34.
22. Lee HM, Kim HS, Kim DJ, et al. Reliability of magnetic resonance imaging in detecting posterior ligament complex injury in thoracolumbar spinal fractures. *Spine (Phila Pa 1976)* 2000;25:2079–84.
23. Blauth M, Bastian L, Knop C, et al. Inter-observer reliability in the classification of thoraco-lumbar spinal injuries. *Orthopade* 1999;28:662–81.
24. Schnake KJ, von Scotti F, Haas NP, et al. Type B injuries of the thoracolumbar spine : Misinterpretations of the integrity of the posterior ligament complex using radiologic diagnostics. *Unfallchirurg* 2008;111:977–84.
25. Nicoll EA. Fractures of the dorso-lumbar spine. *J Bone Joint Surg Br* 1949;31B:376–94.
26. Vaccaro AR, Lim MR, Hurlbert RJ, et al. Surgical decision making for unstable thoracolumbar spine injuries: results of a consensus panel review by the spine trauma study group. *J Spinal Disord Tech* 2006;19:1–10.
27. Patel AA, Vaccaro AR, Albert TJ, et al. The adoption of a new classification system: time-dependent variation in interobserver reliability of the thoracolumbar injury severity score classification system. *Spine (Phila Pa 1976)* 2007;32:E105–10.
28. Raja Rampersaud Y, Fisher C, Wilsey J, et al. Agreement between orthopedic surgeons and neurosurgeons regarding a new algorithm for the treatment of thoracolumbar injuries: a multicenter reliability study. *J Spinal Disord Tech* 2006;19:477–82.
29. Audige L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand* 2004;75:184–94.