The reliability of the AOSpine Thoracolumbar Spine Injury Classification System in children: an international validation study

Andrew Z. Mo[1]
Patricia E. Miller[2]
Javier Pizones[3]
Ilkka Helenius[4]
Michael Ruf[5]
Ron El-Hawary[6]
Rafael Garcia de Oliveira[7]
Dror Ovadia[8]
Noriaki Kawakami[9]
Haemish Crawford[10]
Thierry Odent[11]
Muharrem Yazici[12]
Michael B. Johnson[13]
Firoz Miyanji[14]
Daniel J. Hedequist[2]

Abstract

Purpose To evaluate the AOSpine Thoracolumbar Spine Injury Classification System and if it is reliable and reproducible when applied to the paediatric population globally.

Methods A total of 12 paediatric orthopaedic surgeons were asked to review MRI and CT imaging of 25 paediatric patients with thoracolumbar spine traumatic injuries, in order to determine the classification of the lesions observed. The evaluators classified injuries into primary categories: A, B and C. Interobserver reliability was assessed for the initial reading by Fleiss's kappa coefficient (k_F) along with 95% confidence intervals (CI). For A and B type injuries, sub-classification was conducted including A0-A4 and B1-B2 subtypes. Interobserver reliability across subclasses was assessed using Krippendorff's alpha (α_k) along with bootstrapped 95% CIs. A second round of classification was performed one-month later. Intraobserver reproducibility was assessed for the primary classifications using Fleiss's kappa and sub-classification reproducibility was assessed by Krippendorff's alpha (α_k) along with 95% CIs.

Results In total, 25 cases were read for a total of 300 initial and 300 repeated evaluations. Adjusted interobserver reliability was almost perfect ($kF = 0.74$; 95% CI 0.71 to 0.78) across all observers. Sub-classification reliability was substantial ($\alpha_k = 0.67$; 95% CI 0.51 to 0.81), Adjusted intraobserver reproducibility was almost perfect ($kF = 0.91$; 95% CI 0.83 to 0.99) for both primary classifications and for sub-classifications ($\alpha_k = 0.88$; 95% CI 0.83 to 0.93).

Conclusion The inter- and intraobserver reliability for the AOSpine Thoracolumbar Spine Injury Classification System was high amongst paediatric orthopaedic surgeons. The AOSpine Thoracolumbar Spine Injury Classification System is a promising option as a uniform fracture classification in children.

Level of Evidence: III

[1] Department of Orthopaedic Surgery, Georgetown University Medical Center, Washington, District of Columbia, United States
[2] Department of Orthopaedic Surgery, Boston Children's Hospital, Boston, Massachusetts, United States
[3] Spine Unit, Hospital Universitario La Paz, Madrid, Spain
[4] Professor and Chairman, Department of Orthopaedics and Traumatology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[5] Center for Spinal Surgery, Orthopedics, and Traumatology, SRH Klinikum Karlsbad-Langensteinbach, Karlsbad, Germany
[6] IWK Health Centre, Halifax, Nova Scotia, Canada
[7] Sarah Hospital, Brasília, Brazil
[8] Dana Dwek Children's Hospital, Tel Aviv Medical Center, Tel Aviv, Israel
[9] Director of Spine & Scoliosis Center, Department of Orthopedic Surgery, Ichinomiyanishi Hospital, Ichinomiya, Japan
[10] Paediatric Orthopaedic Surgeon, Starship Children's Hospital, Auckland, New Zealand
[11] Service de Chirurgie Orthopédique Pédiatrique, Université François Rabelais de Tours, PRES Centre-Val de Loire Université, Hôpital Gatien-de-Clocheville, Tours, France
[12] Hacettepe University, Faculty of Medicine, Orthopaedics, Ankara, Turkey
[13] Royal Children's Hospital, Melbourne, Victoria, Australia
[14] Department of Orthopedics, British Columbia Children's Hospital, Pediatric Orthopedics and Spine Surgery, Vancouver, British Columbia, Canada

Correspondence should be sent to Daniel J. Hedequist, Department of Orthopedic Surgery, Boston Children's Hospital, 300 Longwood Ave, Boston, MA 02115 United States.
E-mail: daniel.hedequist@childrens.harvard.edu

Introduction

The AOSpine Thoracolumbar Spine Injury Classification System is the newest widespread classification system for thoracolumbar trauma injuries, improving on previous classifications that evolved from simple morphological classifications to more complex systems.[1-3] Within the paediatric population there does not exist a dedicated, universal classification. The AOSpine Thoracolumbar Spine Injury Classification System incorporates fracture morphology (injury mechanism), evidence of posterior ligamentous integrity, neurological status of the patient, as well as patient specific modifiers to classify injuries. It incorporates and expands on elements of the Magerl classification system, Denis classification system and Thoracolumbar Injury Classification System (TLICS).[4,5] TLICS was a milestone, unifying many previous classification systems. It was also one of the first in providing treatment guidelines, although this provided only definitive recommendations for clear cases. TLICS has been validated in several studies as a reliable classification system in the paediatric population.[6-8] In a previous study, TLICS and the AOSpine Thoracolumbar Spine Injury Classification system has been found to corelate when assessed in the peadiatric population.[9]

The lead authors of this study previously assessed the inter- and intraobserver reliability of the AOSpine Thoracolumbar Spine Injury Classification System in children with a single country cohort of observers.[10] The aim of this study was to confirm the interobserver reliability and intraobserver reproducibility of the AOSpine Thoracolumbar Spine Injury Classification System with an international cohort of observers when applied to the paediatric population on a global level, representing diverse patient populations and reviewers with diverse training backgrounds.

Materials and methods

A retrospective institutional review was performed using Boston Children's Hospital's internal trauma database.

Institutional review board approval for this study was obtained. Inclusion criteria included patients under the age of 18 years who had been treated operatively for a thoracolumbar fracture between 2006 and 2016 and had available CT scans and MRI.

Imaging records were collected and de-identified. Each patient case included plain film radiographs, CT and MRI. CT and MRI were exported as cine clips (SYNPASE PACS; FUJIFILM Medical Systems USA Inc, Stamford, Connecticut). Each patient case was uploaded to an online survey interface (Google Forms; Alphabet Inc, Mountain View, California), which were divided into three forms each consisting of sets of patient cases to be reviewed and classified.

Radiographic assessment of thoracolumba spinal injuries using the AOSpine Thoracolumbar Spine Injury Classification System (A0-4, B1-2 or C) was conducted by 12 paediatric orthopaedic surgeons (JP, IH, MR, REH, RGO, DO, NK, HC, TO, MY, MJ, and FM) with experience in paediatric spine surgery and paediatric spine trauma. Reviewers were selected from international academic centres and include paediatric surgeons from Australia, Brazil, Canada, Finland, France, Germany, Israel, Japan, New Zealand and Turkey. The 12 evaluators classified injuries into three primary categories: A, B or C. Injury morphology was classified as an A injury (compression), B injury (distraction) or C injury (translation). For each patient case, if multiple injuries were present, the most severe injury was recorded and classified. Type A fractures were graded in increasing severity as follows: A0 (simple), A1 (compression), A2 (pincer), A3 (burst involving one endplate) (Fig. 1) and A4 (burst involving both endplates). Type B fractures include classic bony chance (B1), failure of the posterior tension band such as horizontal fracture lines through the posterior elements or evidence of posterior ligamentous disruption (B2) (Fig. 2) and hyperextension injuries (B3). Type C fractures/injuries demonstrate dissociation between cranial and caudal segments (Fig. 3).

In total, 25 patient cases met inclusion criteria and were distributed into three forms consisting of eight, eight and nine sets of patient imaging. An initial test form was sent to familiarize observers with the interface, utilizing a sample case and imaging. The observers were each provided a poster illustration of the classification system as well as a video tutorial. In the initial review, each reviewer completed the three forms. After a one-month interval from the initial review, the same patient cases were randomized and distributed into three new sets and redistributed for repeat review (Fig. 4).

Statistical analysis

Intraobserver reproducibility was assessed for the primary classifications using Fleiss's kappa and sub-classification reproducibility was assessed by Krippendorff's alpha (α_k)

Fig. 1 CT sagittal image demonstrating A3 injury of the L1 vertebra.

along with 95% CIs. Interobserver reliability was assessed for the initial reading across all 12 observers by Fleiss's kappa coefficient (k_F) along with 95% confidence intervals (CIs). For A and B type injuries, sub-classification was conducted including A0-A4 and B1-B2 subtypes. Interobserver reliability across subclasses was assessed using Krippendorff's alpha (α_k) along with bootstrapped 95% CIs. Interpretations for reliability estimates were based on Landis and Koch: 0 to 0.2, slight; 0.2 to 0.4, fair; 0.4 to 0.6, moderate; 0.6 to 0.8, substantial; and > 0.8, almost perfect agreement.[11]

Results

In all, 25 patients met inclusion criteria. Demographics are included in Table 1. Utilizing the AOSpine Thoracolumbar Spine Injury Classification System: six patients had type A injuries, 15 patients had type B injuries and four patients had C injuries. Subtype representation in the patient cases are available in Table 2.

Fig. 2 MRI STIR (short T1//tau inversion recovery) sagittal cut demonstrating L2-3 B2 and L2 A3 injury with posterior ligamentous complex disruption.

Interobserver reliability

Adjusted interobserver agreement for primary classifications (A, B and C) was 74% (k_F = 0.74; 95% CI 0.71 to 0.78) on the initial read, suggesting almost perfect agreement across 12 observers, with exact, unadjusted interobserver agreement occurring in 44% (11/25) of cases (Table 3). Adjusted sub-classification agreement was 68% (α_K = 0.67; 95% CI 0.51 to 0.81), indicating substantial agreement, with exact, unadjusted occurring in 16% (4/25) of cases (Table 3).

Intraobserver reproducibility

Adjusted intraobserver agreement for primary classifications (A, B, C) was 91% (k_F = 0.91; 95% CI 0.83 to 0.99), indi-

Fig. 3 CT sagittal cut demonstrating L2-3 C injury with L2-3 translation.

cating almost perfect agreement, with exact, unadjusted intraobserver agreement occurring in 94% (282/300) of ratings across all 12 observers (Table 4). Interobserver agreement for each observer ranged from 0.81 to 1.00. Adjusted intraobserver agreement for sub-classifications was 88% (α_k = 0.88; 95% CI 0.83 to 0.93), suggesting almost perfect agreement, with exact, unadjusted intraobserver agreement occurring in 81% (243/300) of cases

across all 12 observers. Sub-classification interobserver agreement for each observer ranged from 0.77 to 0.95 (Table 4).

Discussion

Thoracolumbar spine trauma classifications have evolved and expanded significantly throughout the development of spine surgery.[1,4] Several studies and classifications have been proposed since Böhler in 1929.[12] As previously noted, TLICs was a pivotal step, uniting key aspects of previous classifications. By incorporating values on fracture morphology, integrity of the posterior ligamentous complex integrity, and neurological status, it also attempted to provide treatment guidelines via a point system.[3] While a step in the right direction, treatment recommendations were otherwise unclear when the case was not clearly operative or nonoperative.

While exhaustive efforts have been put forth to create and validate classifications for thoracolumbar injuries in the adult spine, there is no uniform classification system in the paediatric population. Additionally, there are no operative guidelines for children with spinal trauma. Paediatric thoracolumbar fractures are commonly grouped by morphology into compression fractures, burst fractures, chance injuries and injuries with translation. Many translate the principles of the TLICS to guide treatment in paediatric spine injuries. There are several studies that have validated the use of TLICs in the paediatric population.[6-8]

The AOSpine Thoracolumbar Spine Injury Classification System aims to improve and adapt key accepts of the classifications that came before it. Building off of the AO Magerl classification and TLICS, it is the result of the work of surgeons in the AOSpine Classification group in the AOSpine Knowledge Forum. The classification ranges from simple avulsion fractures of the spine (A0) in a patient with no neurological injury to severe translational injuries (type C) with complete neurological loss.[2,13,14] The classification integrates fracture morphology, integrity of the posterior ligamentous integrity and the neurological status of the patient.

The status of the PLC is a key component of this classification. Severity of PLC injury is believed to correlate with subsequent potential for instability, deformity and neurological compromise.[15,16] CT scans can be used to indirectly evaluate PLC integrity; MRI can add accuracy to the assessment.[17]

The AOSpine Thoracolumbar Spine Injury Classification System has previously been found to have good inter- and intraobserver reliability and reproducibility in the adult population.[14,18,19] When evaluating the AOSpine Thoracolumbar Spine Injury Classification System in children,
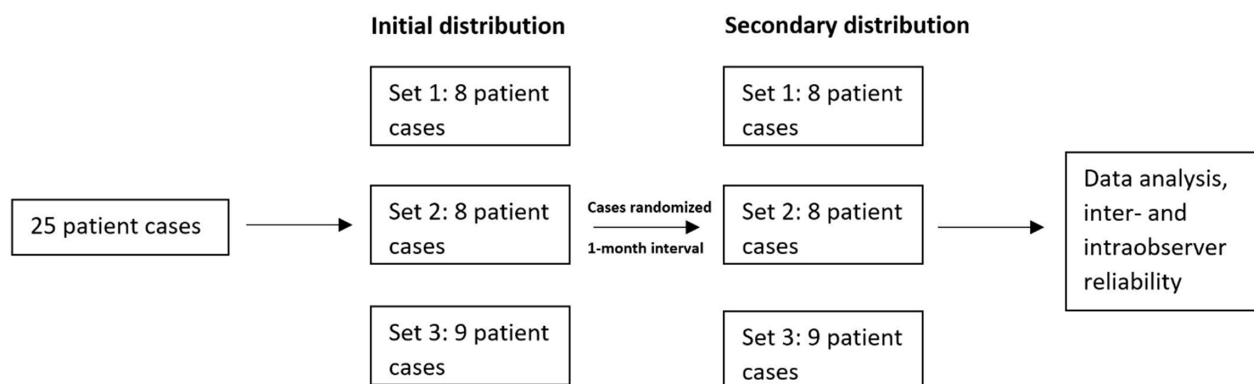
Initial distribution

Secondary distribution

25 patient cases → Set 1: 8 patient cases

Set 2: 8 patient cases — Cases randomized 1-month interval → Set 2: 8 patient cases

Set 3: 9 patient cases

Set 1: 8 patient cases

Set 3: 9 patient cases → Data analysis, inter- and intraobserver reliability

Fig. 4 Study design.

Table 1 Patient, injury, and surgical characteristics (n = 25)

Characteristic	Mean/frequency
Mean age, yrs (SD)	13.6 (3.61)
Mechanism of injury, n (%)	
Motor vehicle accident	15 (60)
Fall	7 (28)
Sports related	3 (12)
Procedure type, n (%)	
Posterior	25 (100)
Injury classification	
AOSpine, n (%)	
A	6 (24)
B	14 (56)
C	5 (20)

Table 2 Distribution of thoracolumbar injuries (n = 300) for each read

AO classification	Read 1, n (%)	Read 2, n (%)
A1	7 (3)	6 (3)
A2	5 (2)	10 (4)
A3	55 (24)	49 (22)
A4	56 (25)	56 (25)
A5	5 (2)	5 (2)
B1	37 (16)	37 (16)
B2	69 (31)	70 (31)
B3	1 (0)	3 (1)
C	65 (29)	64 (28)

using a single country cohort of observers and the same patient series as this study, the results demonstrated high interobserver reliability of 82% (k_F = 0.82; 95% CI 0.77 to 0.87) and intraobserver reproducibility of 81% (k_F = 0.81; 95% CI 0.71 to 0.90).[10] The results of this study using an international group of observers are comparable, with interobserver reliability of 74% (kF = 0.74; 95% CI 0.71 to 0.78) and intraobserver reliability of 91% (kF = 0.91; 95% CI 0.83 to 0.99). Given the diverse group of global reviewers with different patient populations and educational backgrounds, the results of this study further strengthen the applicability of the AOSpine Thoracolumbar Spine Injury Classification System in the paediatric population.

Table 3 Interobserver reliability across 12 observers

Classification	Coefficient	95% confidence interval
Primary classifications (A, B and C)	k_F	
Read 1	0.74	0.71 to 0.78
Read 2	0.73	0.69 to 0.76
All reads	0.74	0.71 to 0.76
Sub-classifications (A0-A4, B1-B2, C)	α_k	
Read 1	0.68	0.51 to 0.81
Read 2	0.65	0.45 to 0.80
All reads	0.67	0.53 to 0.77

α_k, Krippendorff's alpha coefficient; k_F, Fleiss's kappa coefficient

This study was limited by the decreased representation of certain subtypes, specifically a lack of cases demonstrating A1, A2, A3 and B3 subtypes. Additionally, this study focused on fracture pattern without consideration of neurological status, a core component of the AOSpine Thoracolumbar Spine Injury Classification System.[18] Due to the retrospective nature of this study, information on neurological status was limited. Additionally, within this study all patients had a CT and MRI scan, which may not always be available. Finally, given the age range of the patients in this cohort, our study could not account for the anatomic differences of varying skeletal maturity.

Conclusion

The AOSpine Thoracolumbar Spine Injury Classification System shows considerable promise in the paediatric population, potentially obviating the need to create a new paediatric classification *de novo*. This would save both time and resources while providing a useful tool for consistent physician communication. Similar to a previous single country cohort of observers, the results of this study incorporating an global cohort of observers show high interobserver reliability and interobserver reproducibility when applying the AOSpine Thoracolumbar Spine Injury

Table 4 Intraobserver reliability across all 12 observers and by individual observer

Classification	Coefficient	95% confidence interval	Percentage agreement
Primary classifications (A, B and C)	k_F		
All raters	0.91	0.83 to 0.99	94
Rater 1	0.87	0.59 to 1	92
Rater 2	0.93	0.64 to 1	96
Rater 3	0.94	0.65 to 1	96
Rater 4	0.94	0.66 to 1	96
Rater 5	0.93	0.64 to 1	96
Rater 6	0.94	0.65 to 1	96
Rater 7	1.00	0.71 to 1	100
Rater 8	0.81	0.53 to 1	88
Rater 9	0.87	0.59 to 1	92
Rater 10	0.87	0.59 to 1	92
Rater 11	0.82	0.54 to 1	88
Rater 12	1.00	0.71 to 1	100
Sub-classifications (A1-A5, B1-B2, C)	α_k		
All raters	0.88	0.83 to 0.93	81
Rater 1	0.81	0.50 to 1.00	88
Rater 2	0.85	0.55 to 1.00	80
Rater 3	0.93	0.79 to 0.99	80
Rater 4	0.88	0.60 to 0.99	80
Rater 5	0.90	0.69 to0.99	80
Rater 6	0.94	0.82 to 1.00	84
Rater 7	0.94	0.84 to 0.99	80
Rater 8	0.77	0.43 to 0.97	76
Rater 9	0.93	0.80 to 1.00	84
Rater 10	0.87	0.62 to 0.99	80
Rater 11	0.83	0.60 to 0.98	76
Rater 12	0.95	0.83 to 1.00	88

α_k, Krippendorff's alpha coefficient; kF, Fleiss's kappa coefficient

Classification System to the paediatric population.[10] Additionally, these results are strengthened with the diverse group comprising the study observers. Further studies are needed to determine the transferability of the AOSpine Thoracolumbar Spine Injury Classification System in the other anatomical regions of the spine, which are currently underway.

COMPLIANCE WITH ETHICAL STANDARDS

FUNDING STATEMENT

OA LICENCE TEXT

ETHICAL STATEMENT
Ethical approval: No human participants and/or animals were involved. This article does not contain any studies with animals performed by the author. Institution Review Board approval was obtained for this study.

Informed consent: No informed consent was necessary in this study, as all patient data was de-identified prior to collection.

ICMJE CONFLICT OF INTEREST STATEMENT
The authors indicated that no external funding was received for any aspect of this work. None of the authors have any conflicts of interest to disclose.

AUTHOR CONTRIBUTIONS
AZM: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

PEM: Analysis and interpretation of data, Drafting and revision of the manuscript.

JP: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

IH: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

MR: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

REH: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

RGdO: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

DO: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

NK: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

HC: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

TO: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

MY: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

MBJ: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

FM: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

DJH: Study design, Analysis and interpretation of data, Drafting and revision of the manuscript.

REFERENCES

1. Sethi MK, Schoenfeld AJ, Bono CM, Harris MB. The evolution of thoracolumbar injury classification systems. *Spine J.* 2009;9:780-788.

2. Schroeder GD, Harrop JS, Vaccaro AR. Thoracolumbar Trauma Classification. *Neurosurg Clin N Am.* 2017;28:23-29.

3. Patel AA, Vaccaro AR. Thoracolumbar spine trauma classification. *J Am Acad Orthop Surg.* 2010;18:63-71.

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

5. 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.* 2005;30:2325-2333.

6. Savage JW, Moore TA, Arnold PM, et al. The reliability and validity of the Thoracolumbar Injury Classification System in Pediatric Spine Trauma. *Spine.* 2015;40:E1014-E1018.

7. Sellin JN, Steele WJ III, Simpson L, et al. Multicenter retrospective evaluation of the validity of the Thoracolumbar Injury Classification and Severity Score system in children. *J Neurosurg Pediatr.* 2016;18:164-170.

8. Dawkins RL, Miller JH, Ramadan OI, et al. Thoracolumbar Injury Classification and Severity Score in children: a reliability study. *J Neurosurg Pediatr.* 2018;21:284-291.

9. Mo AZ, Miller PE, Troy MJ, Rademacher ES, Hedequist DJ. The AOSpine thoracolumbar spine injury classification system: a comparative study with the thoracolumbar injury classification system and severity score in children. *OTA Int.* 2019;2:e036.

10. Mo AZ, Miller PE, Glotzbecker MP, et al. The reliability of the AOSpine Thoracolumbar Classification System in children: results of a multicenter study. *J Pediatr Orthop.* 2020;40:e352-e356.

11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.

12. Böhler L. *Die Techniek de Knochenbruchbehandlung im Frieden und im Kriege.* Vienna, Austria: Verlag von Wilhelm Maudrich, 1929.

13. Schroeder GD, Vaccaro AR, Kepler CK, et al. Establishing the injury severity of thoracolumbar trauma: confirmation of the hierarchical structure of the AOSpine Thoracolumbar Spine Injury Classification System. *Spine.* 2015;40:E498-E503.

14. Urrutia J, Zamora T, Yurac R, et al. An independent interobserver reliability and intraobserver reproducibility evaluation of the new AOSpine Thoracolumbar Spine Injury Classification System. *Spine.* 2015;40:E54-E58.

15. Wu CC, Jin HM, Yan YZ, et al. Biomechanical role of the thoracolumbar ligaments of the posterior ligamentous complex: a finite element study. *World Neurosurg.* 2018;112:e125-e133.

16. Merter A, Karaca MO, Yazar T. Biomechanical effects of sequential resection of the posterior ligamentous complex on intradiscal pressure and resistance to compression forces. *Acta Orthop Traumatol Turc.* 2019;53:502-506.

17. Pizones J, Sánchez-Mariscal F, Zúñiga L, Álvarez P, Izquierdo E. Prospective analysis of magnetic resonance imaging accuracy in diagnosing traumatic injuries of the posterior ligamentous complex of the thoracolumbar spine. *Spine.* 2013;38:745-751.

18. Vaccaro AR, Oner C, Kepler CK, et al. AOSpine thoracolumbar spine injury classification system: fracture description, neurological status, and key modifiers. *Spine.* 2013;38:2028-2037.

19. Kepler CK, Vaccaro AR, Koerner JD, et al. Reliability analysis of the AOSpine thoracolumbar spine injury classification system by a worldwide group of naïve spinal surgeons. *Eur Spine J.* 2016;25:1082-1086.