Korean Journal of Radiology

MRI Inter-Reader and Intra-Reader Reliabilities for Assessing Injury Morphology and Posterior Ligamentous Complex Integrity of the Spine According to the Thoracolumbar Injury Classification System and Severity Score

Guen Young Lee, MD[1], Joon Woo Lee, MD, PhD[1, 2], Seung Woo Choi, MD[1], Hyun Jin Lim, MD[1],
Hye Young Sun, MD[1], Yusuhn Kang, MD[1], Jee Won Chai, MD[3], Sujin Kim, MD[3], Heung Sik Kang, MD[1, 2]

[1]Department of Radiology, Seoul National University Bundang Hospital, Seongnam 463-707, Korea; [2]Department of Radiology and Institute of Radiation Medicine, Seoul National University College of Medicine, Seoul 110-744, Korea; [3]Department of Radiology, SMG-SNU Boramae Medical Center, Seoul 156-707, Korea

Objective: To evaluate spine magnetic resonance imaging (MRI) inter-reader and intra-reader reliabilities using the thoracolumbar injury classification system and severity score (TLICS) and to analyze the effects of reader experience on reliability and the possible reasons for discordant interpretations.

Materials and Methods: Six radiologists (two senior, two junior radiologists, and two residents) independently scored 100 MRI examinations of thoracolumbar spine injuries to assess injury morphology and posterior ligamentous complex (PLC) integrity according to the TLICS. Inter-reader and intra-reader agreements were determined and analyzed according to the number of years of radiologist experience.

Results: Inter-reader agreement between the six readers was moderate (k = 0.538 for the first and 0.537 for the second review) for injury morphology and fair to moderate (k = 0.440 for the first and 0.389 for the second review) for PLC integrity. No significant difference in inter-reader agreement was observed according to the number of years of radiologist experience. Intra-reader agreements showed a wide range (k = 0.538–0.822 for injury morphology and 0.423–0.616 for PLC integrity). Agreement was achieved in 44 for the first and 45 for the second review about injury morphology, as well as in 41 for the first and 38 for the second review of PLC integrity. A positive correlation was detected between injury morphology score and PLC integrity.

Conclusion: The reliability of MRI for assessing thoracolumbar spinal injuries according to the TLICS was moderate for injury morphology and fair to moderate for PLC integrity, which may not be influenced by radiologist' experience.

Index terms: Thoracolumbar; Trauma; MRI; Reliability; Kappa; Radiologist

INTRODUCTION

Thoracolumbar spine trauma is very common and accounts for 75–90% of all spine fractures (1, 2). Various classification systems have been proposed and revised extensively based on anatomic location (e.g., the Denis three-column system) or injury mechanism (e.g., American Orthopedic [AO] classification), but these have limitations in that they cannot explain the patient's neurological state, they overlook the importance of posterior ligamentous complex integrity (PLC) as a prognostic factor, or show poor

inter-reader or intra-reader reliabilities (3-11). To overcome these shortcomings, the Spine Trauma Study Group (STSG) introduced a new thoracolumbar injury classification system, specifically the thoracolumbar injury classification system and severity score (TLICS), in which the type of injury morphology, integrity of PLC, and neurological status are scored (12). Although reliability of the TLICS is only fair in early studies (13-15), many authors have demonstrated better reliability of the TLICS to shed light on its usefulness in daily practice (16-24). However, these studies were carried out by clinicians, such as neurosurgeons, orthopedic surgeons, or trauma surgeons, who have the patient's clinical information, whereas radiologists actually read the majority of images in many hospitals. Therefore, we were curious about the reliability of TLICS among radiologists, based on magnetic resonance imaging (MRI), as MRI may be the most useful imaging modality for evaluating spine trauma, particularly PLC (25, 26). No study has reported on the TLICS from the radiologist's perspective; thus, the aim of this retrospective study was to evaluate inter-reader and intra-reader reliabilities of MRI in the TLICS for assessing thoracolumbar spinal injuries and to analyze the effect of reader experience on the reliability and the possible reasons for discordant interpretations.

MATERIALS AND METHODS

Patients

This retrospective study was approved by the Institutional Review Board of our hospital, and informed consent was waived. Sample size was estimated as follows. If the expected lower boundary for a kappa one-sided 95% confidence interval (CI) was 0.35 and the expected preliminary kappa value was 0.5, a minimum of 100 subjects was required for this study of inter-reader agreement by six raters, with prevalence of 0.1. We estimated sample size

using the kappaSize library statistical program in R-project (R Core Team [2012]. R: A language and environment for statistical computing; R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/) (27). We included patients who had visited the emergency room of our hospital for suspected thoracolumbar traumatic injury from July 2010 to July 2011 and underwent thoracolumbar MRI within 24 hours after their trauma. Exclusion criteria were 1) history of thoracolumbar spine surgical history and 2) poor image quality. After searching our picture archiving and communication system, 100 consecutive thoracolumbar spine MRI examinations were enrolled in the study. There were 45 males and 55 females with a mean age of 57.6 years (range, 11–89 years). Among the 100 patients, 79 received conservative management, including vertebroplasty and 21 underwent surgery.

MRI Examinations

Of the 100 MRI examinations, 15 performed were outside of the hospital before visiting our emergency room. Eighty-five MRI examinations were performed in our hospital, of which 50 involved contrast enhancement (gadodiamide [Omniscan]; GE Healthcare, Princeton, NJ, USA). Seventy-five patients were scanned in the lumbar spine area and 22 had their thoracolumbar spine investigated by MRI. Three patients underwent whole spine MRI scanning from the skull base to the sacrococcygeal junction on the sagittal plane, and contrast enhancement was used in two. T2-weighted short time inversion recovery (STIR) sagittal scanning was also performed in 89 patients. Ninety-eight MRI examinations including all 15 outside hospital examinations were performed using 1.5 T MRI (83 examinations in our hospital with an Integra 1.5 T, Philips, Best, the Netherlands) and two were done using a 3 T MRI (Achieva 3T, Philips). The MRI parameters are listed in Table 1.

Table 1. MR Parameters of Thoracolumbar Spines

	TR (ms)	TE (ms)	FA	ETL	NEX	Matrix	Slice Thickness (mm)	Slice Gap (mm)
T2W axial	2518–6805	100–120	90°	16–29	2–3	216 x 174–512 x 512	4–8	4–10
T2W sagittal	2769–4182	100–120	90°	16–43	2	352 x 297–512 x 512	3–4	3–4
STIR sagittal	2330–2417	80–100	90°	8–36	2	268 x 200–512 x 512	3–4	3–4
T1W sagittal	408–569	9–17	90°	3–8	2	444 x 220–512 x 512	3–4	3–4
Enhanced T1WFS sagittal	400–633	7–20	90°	3–4	2–3	224 x 201–512 x 512	3–4	3–4
Enhanced T1WFS axial	492–600	7–17	90°	4–10	2–3	208 x 165–512 x 512	4–8	4–10

ETL = echo train length, FA = flip angle, NEX = number of excitation, STIR = short tau inversion recovery, TE = echo time, TR = repetition time, T1W = T1-weighted, T1WFS = T1-weighted fat saturated, T2W = T2-weighted

Image Interpretation

All MRI images were reviewed independently by six radiologists (two senior radiologists; one specializing in spine radiology with 11 years experience and one with 4 years experience), two junior radiologists (both specializing in spine radiology with 1 year experience), and two residents (one fourth-year and one third-year). Before the image evaluation, all six radiologists were asked to read two articles about TLICS (12, 25) to understand the use of TLICS in terms of reading spine MRI examinations, without a consensus meeting. The six radiologists evaluated the 100 MRI images independently and scored injury morphology using four scales of compression fracture, burst fracture, translation/rotation injury, and distraction injury, as well as PLC integrity using three 3 scales of intact, suspected/indeterminate, and injured based on the TLICS. All readers re-evaluated the images 6 months later to determine intra-reader reliability.

Statistical Analysis

Inter-reader agreement between the six readers on the TLICS scores related to injury morphology and PLC integrity was calculated using the Fleiss kappa (k) value. Fleiss' k was used instead of Cohen's k, as it allows comparing multiple readers using multiple scoring, as required in this study (28). The 95% CIs for the k statistic were calculated using the bootstrap method. Inter-reader agreement for the TLICS scores between two readers, such as the two senior radiologists, the two junior radiologists, and the two residents, as well as intra-reader agreement were estimated using Cohen's k. The k value was defined as slight (0–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect (0.81–1.00). The statistical

analysis was performed using Stata 10.0 (StataCorp, College Station, TX, USA). The relationship between injury morphology and PLC integrity was evaluated using Spearman's coefficient analysis and the treatment option (non-surgical vs. surgical treatment) according to the total injury morphology and PLC integrity scores of each reader using the chi-square test and PASW ver. 17.0 software (SPSS Inc., Chicago, IL, USA). The chi-square test was used to compare the TLICS scores for injury morphology and PLC integrity between patients who were treated operatively and non-operatively. A p value < 0.05 was considered significant.

RESULTS

Inter-reader agreement between the six readers was moderate for the TLICS score for injury morphology (k = 0.538 for the first and 0.537 for the second review) and fair to moderate for PLC integrity (k = 0.440 for the first and 0.389 for the second review). After grouping the readers according to experience, the inter-reader agreement between two readers was spread widely, such as moderate to substantial agreement for injury morphology (k = 0.509–0.664) and fair to substantial for PLC integrity (k = 0.349–0.618) regardless of the radiologist's years of experience (Table 2). No difference in inter-reader agreement was detected over time. Intra-reader agreements were moderate to almost perfect for injury morphology (k = 0.538–0.822) and moderate to substantial for PLC integrity (k = 0.423–0.646) without a difference according to years of experience (Table 3). The majority of the 100 patients had a compression or burst injury (injury morphology scores 1–2) (Figs. 1, 2). All readers agreed that 44 patients in

Table 2. Inter-Reader Agreement on MR for Evaluation of Injury Morphology and PLC Integrity

	Fleiss or Cohen Kappa (95% CI)	
	1st Review	2nd Review
Injury morphology		
All six readers	0.538 (0.456–0.622)	0.537 (0.494–0.580)
Two senior radiologists	0.566 (0.424–0.705)	0.535 (0.382–0.689)
Two junior radiologists	0.664 (0.517–0.779)	0.644 (0.500–0.788)
Two residents	0.509 (0.358–0.648)	0.610 (0.461–0.758)
PLC integrity		
All six readers	0.440 (0.361–0.521)	0.389 (0.352–0.429)
Two senior radiologists	0.618 (0.471–0.753)	0.349 (0.162–0.535)
Two junior radiologists	0.446 (0.303–0.609)	0.422 (0.249–0.595)
Two residents	0.429 (0.297–0.584)	0.452 (0.288–0.616)

CI = confidence interval, MR = magnetic resonance, PLC = posterior ligamentous complex

the first and 45 in the second reviews had the same injury morphology (Table 4). However, the readers failed to agree on the injury morphology of > 50% of the patients after the first and second reviews, in which the majority had

Table 3. Intra-Reader Agreement of TLICS

	Cohen Kappa (95% CI)
Injury morphology	
Reader A	0.776 (0.662–0.890)
Reader B	0.655 (0.510–0.800)
Reader C	0.558 (0.400–0.685)
Reader D	0.538 (0.396–0.672)
Reader E	0.560 (0.431–0.698)
Reader F	0.822 (0.688–0.913)
PLC integrity	
Reader A	0.646 (0.507–0.785)
Reader B	0.547 (0.350–0.743)
Reader C	0.542 (0.378–0.676)
Reader D	0.434 (0.280–0.587)
Reader E	0.423 (0.279–0.563)
Reader F	0.582 (0.418–0.712)

CI = confidence interval, PLC = posterior ligamentous complex, TLICS = thoracolumbar injury classification system and severity score

compression vs. burst injuries (Table 4). The readers had difficulties distinguishing between burst and distraction injury in nine patients on the first and seven patients on the second review (Fig. 3). Two patients in the first review and six in the second review received various scores for all injury morphology types by the six readers, including the compression, burst, translation/rotation, and distraction injury patterns.

Forty-one patients in the first review and 38 in the second review were regarded as having PLC status by all readers. Among the cases of disagreement, 27 patients in the first review and 29 in the second review were thought to be "intact" or "indeterminate", and nine in the first review and five in the second review were considered "indeterminate" or "injured". Moreover, four patients in the first review and six in the second review were regarded as "intact" or "injured" with disagreement. Unlike injury morphology, in which only two in the first review and six in the second review had various injury morphology scores, 19 cases in the first review and 22 in the second review showed a wide range of PLC integrity scores (Table 4, Figs. 1, 4). Disrupted PLC was the agreed condition by all readers
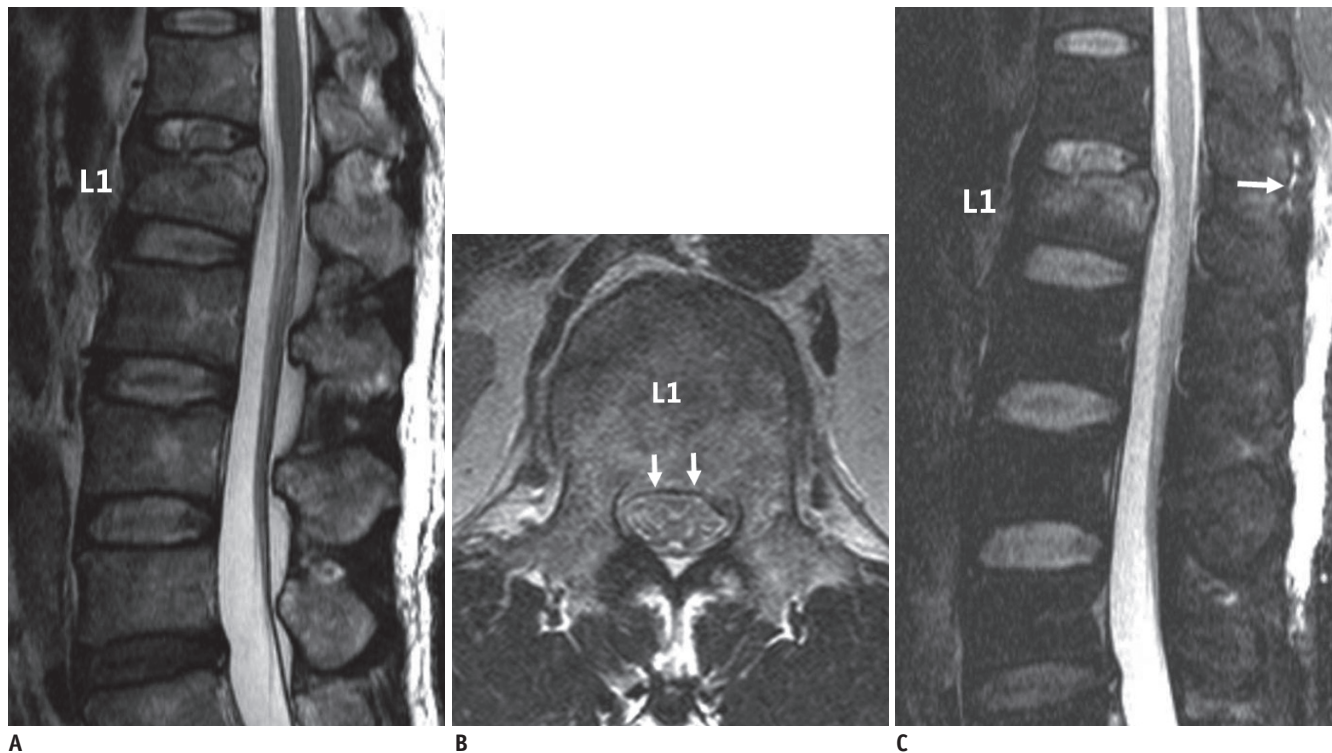

Fig. 1. Case of disagreement about posterior ligamentous complex (PLC) integrity by readers.
49-year-old man showed burst injury of L1 body on T2-weighted sagittal image **(A)** with bulging contour of posterior cortex on T2-weighted axial scan **(B,** arrows), which was agreed by all readers during first and second reviews. Suspicious increase of signal intensity was detected along PLC on T2-weighted short time inversion recovery sagittal image **(C,** arrow), which became disagreement during first ("intact" by one; "indeterminate" by four; "injured" by one) and second reviews ("intact" by one; "indeterminate" by five readers).

for four of 14 patients in the first review and three of 18 in the second review with a suspected distraction injury by any reader.

A positive correlation was detected between injury morphology score and PLC integrity, indicating that a higher injury morphology score resulted in a higher PLC integrity
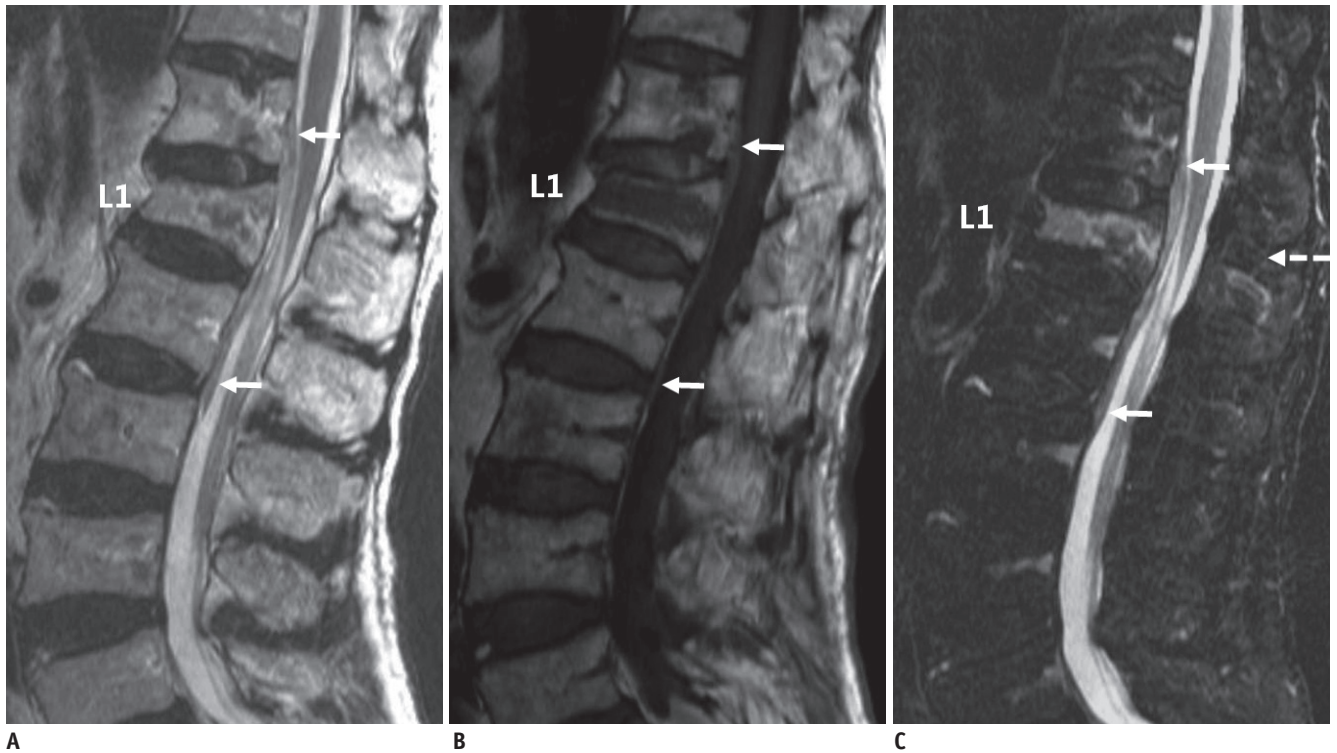

Fig. 2. Case of disagreement about injury morphology and posterior ligamentous complex (PLC) integrity by readers.
89-year-old man suffered traumatic lesion of L1 body with epidural hemorrhage at T12–L2 level (arrows) on T2- **(A)** and T1-weighted sagittal images **(B)**, with minimal bulging of posterior cortex, in which injury morphology became disagreement during first ("compression" by three; "burst" by three readers) and second reviews ("compression" by five; "burst" by one). All readers agreed on intact PLC on T2-weighted short time inversion recovery sagittal image after first review **(C, dashed arrow)** but there was disagreement during second review ("intact" by five; "indeterminate" by one).

Table 4. Agreement Cases for Injury Morphology and PLC Integrity between Six Readers

| Reader	Injury Morphology				PLC Integrity		
	Compression	Burst	Translation/Rotation	Distraction	Intact	Suspected/Indeterminate	Injured
1st review							
A	48	36	0	14	44	34	22
B	32	63	1	4	74	10	16
C	46	41	1	12	69	22	9
D	39	53	1	7	68	11	21
E	59	37	0	4	61	24	15
F	36	58	0	6	48	28	26
All	19	24	0	1	37	0	4
2nd review							
A	40	50	0	10	61	19	20
B	40	55	1	4	81	8	11
C	64	27	0	9	62	20	18
D	55	41	0	4	68	10	22
E	61	36	0	3	66	22	12
F	54	36	0	10	58	24	18
All	28	17	0	0	34	1	3

PLC = posterior ligamentous complex
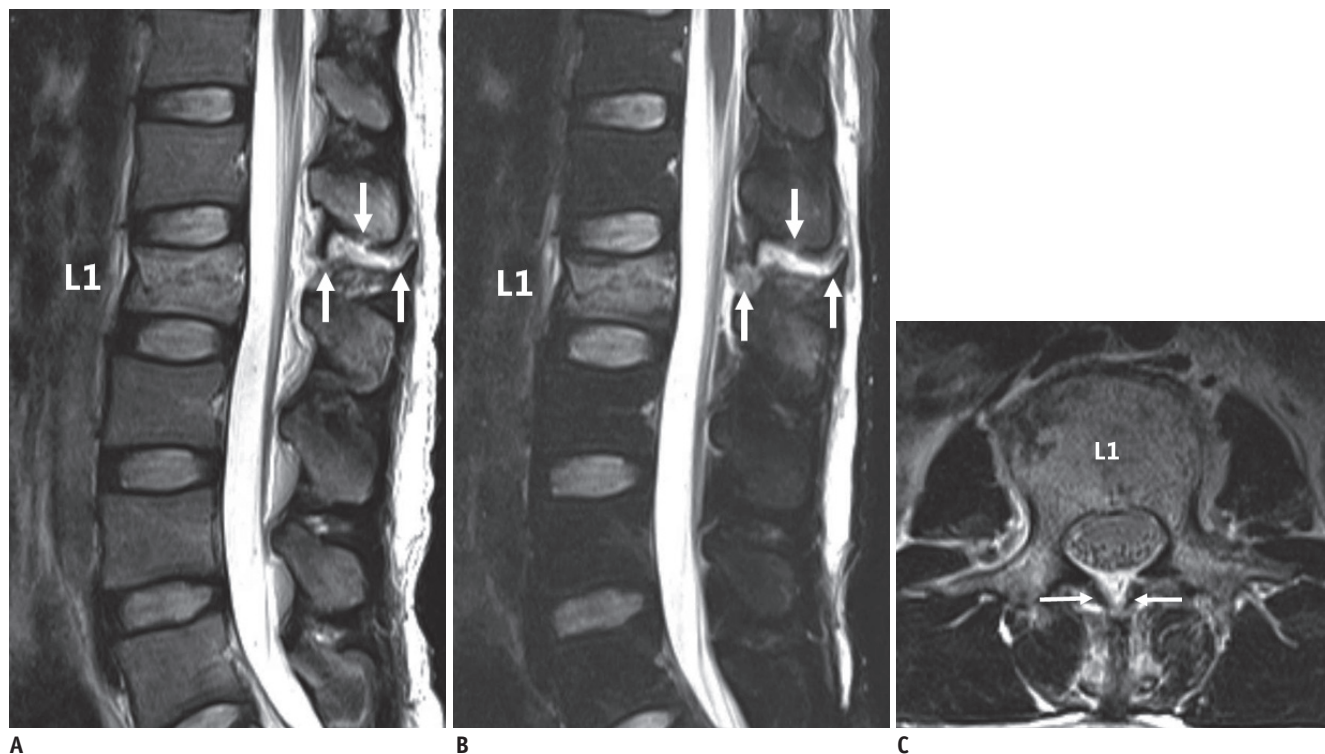
Fig. 3. Case of disagreement on injury morphology by readers.
Posterior ligamentous complex (PLC) disruption was clearly demonstrable and was agreed upon by all six readers during first and second reviews of T2-weighted non-fat saturation **(A)** and short time inversion recovery sagittal images **(B)** and T2-weighted axial scan **(C)** of 33-year-old man (arrows). However, readers failed to agree on injury morphology: "compression" by one, "burst" by one, and "distraction" by four readers during first review and "compression" by one, "burst" by two, and "distraction" by three on second review.

Table 5. Correlation of Scoring between Injury Morphology and PLC Integrity Using TLICS

Reader	Spearman Coefficient	
	1st Review	2nd Review
Reader A	0.516	0.455
Reader B	0.365	0.249
Reader C	0.460	0.614
Reader D	0.635	0.525
Reader E	0.469	0.496
Reader F	0.488	0.503

PLC = posterior ligamentous complex, TLICS = thoracolumbar injury classification system and severity score

score, and vice versa, although there was a variable degree of correlation (Spearman's coefficient, 0.365–0.635 for the first review and 0.249–0.614 for the second) (Table 5).

No difference was observed between non-surgical and surgical treatment according to the total injury morphology and PLC integrity scores (total sum of injury morphology and PLC integrity scores < 5 vs. ≥ 5) by all six readers in the first review. The total score was not different among the four readers on the second review, except two ($p = 0.004$ and 0.012).

The TLICS scores of injury morphology and PLC integrity were significant different between patients who received non-operative or operative management by all six radiologists during the first and second reviews ($p < 0.000$), except for the injury morphology score during the second review by one radiologist with borderline significance ($p = 0.006$).

DISCUSSION

Many classification systems have been proposed, but the Denis and AO classification systems have been widely accepted for spinal instability and recognition of the importance of PLC (29-31). However, these systems have been criticized for their complexity (e.g., AO classification) or their inability to represent future spinal instability. Therefore, the STSG proposed a new TLICS classification system in 2005 (12) and its clinical usefulness has been suggested (22-24). However, reliability studies on the TLICS system have been performed exclusively by neurosurgeons or orthopedic surgeons (13-24), and the reliability of the TLICS has not been evaluated by radiologists.
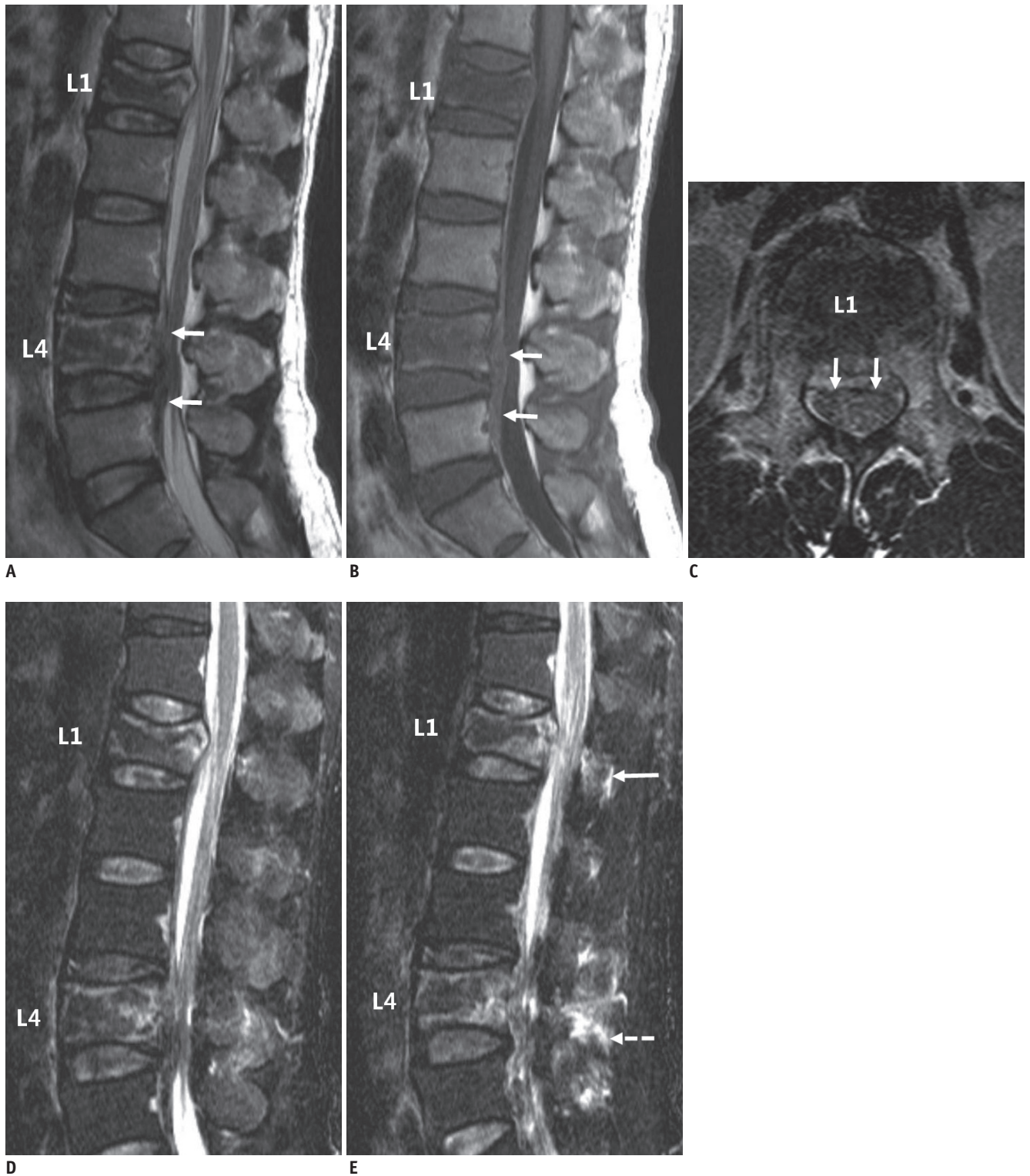
Fig. 4. Case of disagreement about posterior ligamentous complex (PLC) integrity by readers.
51-year-old man demonstrated traumatic lesions at L1 and L4 bodies on T2- **(A)** and T1-weighted sagittal images **(B)** with epidural hemorrhage at L3–5 level (arrows, **A**, **B**). L1 lesion revealed bulging of posterior cortex on T2-weighted axial scan (**C**, arrows), which was agreed to be burst injury by all readers during first review and by five during second review. However, no consensus was reached about PLC integrity on T2-weighted short time inversion recovery sagittal images (**D**, **E**, arrow), which produced variety of scores for PLC intergtity ("intact" by two readers; "indeterminate" by four in first and "intact" by three; "indeterminate" by two; and "injured" by one during second review). In contrast, PLC integrity (**E**, dashed arrow) of L4 body lesion was not evaluated using thoracolumbar injury classification system and was originally classified as most severe injury level.

Inter-reader reliability for injury morphology and PLC integrity of TLICS was only fair in several initial studies but was moderate to almost perfect for neurologic state and management decisions (13-15), which may have influenced their study design, as all readers were neurosurgeons or orthopedic surgeons and were informed about patient clinical neurological status. Later authors demonstrated better inter-reader reliabilities of moderate to substantial (16-18). In our study, conducted exclusively by radiologists, inter-reader agreement of the six radiologists was moderate for injury morphology and fair to moderate for PLC integrity, which was relatively low, and similar to those previous studies. Intra-reader agreement was also relatively low, particularly for PLC integrity with a moderate to substantial k value.

Some authors have evaluated whether reliability differs over time or according to a clinician's experience (19, 20). However, time-dependent improvement in reliability was not explained, because the readers were different in the second session of the study (19) and the difference in k was too minimal to show a significant difference between junior and senior groups (20). In our study, inter-reader agreement between thee sets of two readers revealed no differences regardless of the radiologist's experience.

A previous study reported some confusion regarding injury morphology (32). All readers agreed about the injury morphology of some patients, but the readers had difficulties distinguishing between compression and burst injuries in many patients. Only one case was suspected of suffering a translation/rotation or distraction injury with a disrupted PLC. This result suggests difficulty distinguishing between translation/rotation or distraction injury with definite disruption of the PLC and facet joint dislocation/subluxation.

The usefulness of MRI, particularly T2-weighted fat-suppressed images, for assessing PLC integrity has been suggested (33-40). However, there is no consensus about which finding is most diagnostic when evaluating PLC integrity (41). Two studies have revealed that the reliability of PLC integrity is poor or moderate (42, 43). In contrast, some authors have questioned the role of MRI for assessing PLC integrity, as there are concerns about over- or understanding PLC disruption using MRI findings alone (44, 45). Our readers could not agree about PLC status in > 50% of patients, which may be due, in part, because the most useful MRI findings to diagnose PLC disruption have not been determined, resulting in the ambiguity of

"indeterminate" PLC status on MRI (46). PLC integrity may tend to be determined as disrupted PLC if there is a fracture or marrow edema in the posterior bony compartment. Moreover, the TLICS system scores only the most severe injured level but some problems may occur clinically in cases with discordant features between injury morphology and PLC integrity.

We found that injury morphology severity and PLC integrity were correlated. The TLICS was designed to focus on PLC integrity; thus, radiologists tend to be hyper aware of PLC status on MRI, resulting in assigning more severe injury morphology in cases of suspicious disruption of the PLC.

However, although the intra-reader and inter-reader agreements were relatively low in this study and there was a concern about over-estimating injury morphology and PLC integrity, which may be dependant on each other, a significant difference was detected in injury morphology and PLC integrity between patients who received non-operative or operative management, suggesting that the TLICS system can be used to clinically determine the surgical option and is well-correlated with clinical manifestations.

Our study included 15 MRIs taken at other hospitals with different MRI parameters including T2-weighted STIR sagittal scans, which were appropriate for evaluating PLC integrity. Thus, including outside MRIs was helpful for the TLICS evaluation in an actual clinical setting.

Several limitations of this study should be mentioned. First, the gold standard for accurately assessing injury morphology and PLC integrity was not defined because only 21 patients underwent surgery. Second, we could not define the diagnostic MRI findings for the injury morphology and PLC status of individual cases. Third, only one case of suspected transitional/rotational injury was included, and the majority of cases were compression/burst injuries. Fourth, we could not compare the long-term clinical outcomes according to the TLICS score to determine its clinical and radiological usefulness. Fifth, heterogeneous MRI parameters were used for outside hospital examinations, and a very small number of MRI examinations were performed using 3 T MRI, which may caused heterogeneity in the enrolled cases. Sixth, we did not exclude elderly patients with underlying osteoporosis.

In conclusion, the reliability of MRI for radiologists, who were blinded to clinical information, such as physical examination findings, to evaluate thoracolumbar spinal injury was moderate for injury morphology and fair to

moderate for PLC integrity and did not differ based on the experience of the radiologist.

REFERENCES

1. Hu R, Mustard CA, Burns C. Epidemiology of incident spinal fracture in a complete population. *Spine (Phila Pa 1976)* 1996;21:492-499
2. Patel AA, Vaccaro AR. Thoracolumbar spine trauma classification. *J Am Acad Orthop Surg* 2010;18:63-71
3. Ferguson RL, Allen BL Jr. A mechanistic classification of thoracolumbar spine fractures. *Clin Orthop Relat Res* 1984;(189):77-88
4. Watson-Jones R. The results of postural reduction of the fractures of the spine. *J Bone Joint Surg Am* 1938;20:567-586
5. Denis F. The three column spine and its significance in the classification of acute thoracolumbar spinal injuries. *Spine (Phila Pa 1976)* 1983;8:817-831
6. 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
7. Mirza SK, Mirza AJ, Chapman JR, Anderson PA. Classifications of thoracic and lumbar fractures: rationale and supporting data. *J Am Acad Orthop Surg* 2002;10:364-377
8. Holdsworth F. Fractures, dislocations, and fracture-dislocations of the spine. *J Bone Joint Surg Am* 1970;52:1534-1551
9. Wood KB, Khanna G, Vaccaro AR, Arnold PM, Harris MB, Mehbod AA. Assessment of two thoracolumbar fracture classification systems as used by multiple surgeons. *J Bone Joint Surg Am* 2005;87:1423-1429
10. Oner FC, Ramos LM, Simmermacher RK, Kingma PT, Diekerhof CH, Dhert WJ, 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;11:235-245
11. Blauth M, Bastian L, Knop C, Lange U, Tusch G. [Inter-observer reliability in the classification of thoraco-lumbar spinal injuries]. *Orthopade* 1999;28:662-681
12. Vaccaro AR, Lehman RA Jr, Hurlbert RJ, Anderson PA, Harris M, Hedlund R, 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;30:2325-2333
13. Vaccaro AR, Zeiller SC, Hulbert RJ, Anderson PA, Harris M, Hedlund R, et al. The thoracolumbar injury severity score: a proposed treatment algorithm. *J Spinal Disord Tech* 2005;18:209-215
14. Vaccaro AR, Baron EM, Sanfilippo J, Jacoby S, Steuve J, Grossman E, et al. Reliability of a novel classification system for thoracolumbar injuries: the Thoracolumbar Injury Severity Score. *Spine (Phila Pa 1976)* 2006;31(11 Suppl):S62-S69; discussion S104
15. Raja Rampersaud Y, Fisher C, Wilsey J, Arnold P, Anand N, Bono CM, 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;19:477-482
16. Whang PG, Vaccaro AR, Poelstra KA, Patel AA, Anderson DG, Albert TJ, 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;32:791-795
17. Ratliff J, Anand N, Vaccaro AR, Lim MR, Lee JY, Arnold P, et al. Regional variability in use of a novel assessment of thoracolumbar spine fractures: United States versus international surgeons. *World J Emerg Surg* 2007;2:24
18. Patel AA, Whang PG, Brodke DS, Agarwal A, Hong J, Fernandez C, et al. Evaluation of two novel thoracolumbar trauma classification systems. *Indian J Orthop* 2007;41:322-326
19. Patel AA, Vaccaro AR, Albert TJ, Hilibrand AS, Harrop JS, Anderson DG, 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-E110
20. Lenarz CJ, Place HM, Lenke LG, Alander DH, Oliver D. Comparative reliability of 3 thoracolumbar fracture classification systems. *J Spinal Disord Tech* 2009;22:422-427
21. Koh YD, Kim DJ, Koh YW. Reliability and Validity of Thoracolumbar Injury Classification and Severity Score (TLICS). *Asian Spine J* 2010;4:109-117
22. Lenarz CJ, Place HM. Evaluation of a new spine classification system, does it accurately predict treatment? *J Spinal Disord Tech* 2010;23:192-196
23. Joaquim AF, Fernandes YB, Cavalcante RA, Fragoso RM, Honorato DC, Patel AA. Evaluation of the thoracolumbar injury classification system in thoracic and lumbar spinal trauma. *Spine (Phila Pa 1976)* 2011;36:33-36
24. Lewkonia P, Paolucci EO, Thomas K. 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)* 2012;37:2161-2167
25. Pizones J, Izquierdo E, Alvarez P, Sánchez-Mariscal F, Zúñiga L, Chimeno P, et al. Impact of magnetic resonance imaging on decision making for thoracolumbar traumatic fracture diagnosis and treatment. *Eur Spine J* 2011;20 Suppl 3:390-396
26. Winklhofer S, Thekkumthala-Sommer M, Schmidt D, Rufibach K, Werner CM, Wanner GA, et al. Magnetic resonance imaging frequently changes classification of acute traumatic thoracolumbar spine injuries. *Skeletal Radiol* 2013;42:779-786
27. Donner A, Rotondi MA. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. *Int J Biostat* 2010;6:Article 31
28. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378-382
29. Patel AA, Dailey A, Brodke DS, Daubs M, Harrop J, Whang PG, et al. Thoracolumbar spine trauma classification: the

Thoracolumbar Injury Classification and Severity Score system and case examples. *J Neurosurg Spine* 2009;10:201-206

30. Lee JY, Vaccaro AR, Lim MR, Oner FC, Hulbert RJ, Hedlund R, et al. Thoracolumbar injury classification and severity score: a new paradigm for the treatment of thoracolumbar spine trauma. *J Orthop Sci* 2005;10:671-675

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

32. Schweitzer KM Jr, Vaccaro AR, Lee JY, Grauer JN; Spine Trauma Study Group. Confusion regarding mechanisms of injury in the setting of thoracolumbar spinal trauma: a survey of The Spine Trauma Study Group (STSG). *J Spinal Disord Tech* 2006;19:528-530

33. McArdle CB, Crofford MJ, Mirfakhraee M, Amparo EG, Calhoun JS. Surface coil MR of spinal trauma: preliminary experience. *AJNR Am J Neuroradiol* 1986;7:885-893

34. Emery SE, Pathria MN, Wilber RG, Masaryk T, Bohlman HH. Magnetic resonance imaging of posttraumatic spinal ligament injury. *J Spinal Disord* 1989;2:229-233

35. Kliewer MA, Gray L, Paver J, Richardson WD, Vogler JB, McElhaney JH, et al. Acute spinal ligament disruption: MR imaging with anatomic correlation. *J Magn Reson Imaging* 1993;3:855-861

36. Petersilge CA, Pathria MN, Emery SE, Masaryk TJ. Thoracolumbar burst fractures: evaluation with MR imaging. *Radiology* 1995;194:49-54

37. Terk MR, Hume-Neal M, Fraipont M, Ahmadi J, Colletti PM. Injury of the posterior ligament complex in patients with acute spinal trauma: evaluation by MR imaging. *AJR Am J Roentgenol* 1997;168:1481-1486

38. Lee HM, Kim HS, Kim DJ, Suk KS, Park JO, Kim NH. Reliability of magnetic resonance imaging in detecting posterior ligament complex injury in thoracolumbar spinal fractures. *Spine (Phila Pa 1976)* 2000;25:2079-2084

39. Haba H, Taneichi H, Kotani Y, Terae S, Abe S, Yoshikawa H, et al. Diagnostic accuracy of magnetic resonance imaging for detecting posterior ligamentous complex injury associated with thoracic and lumbar fractures. *J Neurosurg* 2003;99(1 Suppl):20-26

40. Crosby CG, Even JL, Song Y, Block JJ, Devin CJ. Diagnostic abilities of magnetic resonance imaging in traumatic injury to the posterior ligamentous complex: the effect of years in training. *Spine J* 2011;11:747-753

41. Vaccaro AR, Lee JY, Schweitzer KM Jr, Lim MR, Baron EM, Oner FC, et al. Assessment of injury to the posterior ligamentous complex in thoracolumbar spine trauma. *Spine J* 2006;6:524-528

42. Dai LY, Ding WG, Wang XY, Jiang LS, Jiang SD, Xu HZ. Assessment of ligamentous injury in patients with thoracolumbar burst fractures using MRI. *J Trauma* 2009;66:1610-1615

43. Schweitzer KM, Vaccaro AR, Harrop JS, Hurlbert J, Carrino JA, Rechtine GR, et al. Interrater reliability of identifying indicators of posterior ligamentous complex disruption when plain films are indeterminate in thoracolumbar injuries. *J Orthop Sci* 2007;12:437-442

44. Vaccaro AR, Rihn JA, Saravanja D, Anderson DG, Hilibrand AS, Albert TJ, et al. Injury of the posterior ligamentous complex of the thoracolumbar spine: a prospective evaluation of the diagnostic accuracy of magnetic resonance imaging. *Spine (Phila Pa 1976)* 2009;34:E841-E847

45. Rihn JA, Yang N, Fisher C, Saravanja D, Smith H, Morrison WB, 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;12:391-396

46. Pizones J, Zúñiga L, Sánchez-Mariscal F, Alvarez P, Gómez-Rice A, Izquierdo E. MRI study of post-traumatic incompetence of posterior ligamentous complex: importance of the supraspinous ligament. Prospective study of 74 traumatic fractures. *Eur Spine J* 2012;21:2222-2231