Assessment of Interobserver Variability for CT Scan–Based Evaluation of Posterior Ligament Complex Injury in Thoracolumbar Fractures: An International Multicenter Pilot Study

Anupama Maheswaran, DNB, FRCR[1], Siddharth N. Aiyer, MS, FNB[1], Osama Farouk, MD[2], Mohammad El-Sharkawi, MD[2], Jong-Beom Park, MD, PhD[3], Andrew O'Brien, FRCS[4], Ige Oluwole, MD[5], Yue Wang, MD[6], André de Oliveira Arruda, MD[7], and Shanmuganathan Rajasekaran, PhD[1]

Abstract

Study Design: A multicenter, pilot study, for radiological assessment of thoracolumbar spine fractures was performed with the participation of 7 centers in Africa, Europe, Asia, and South America as a part of the AO Foundation network.

Objectives: To determine the interobserver variability for computed tomography (CT) scan–based evaluation of posterior ligament complex (PLC) injury in thoracolumbar fractures.

Methods: Forty-two observers including 1 principal investigator at each participating center performed variability assessment. Each center contributed toward a total of 91 patient images with A3 or A4 thoracolumbar burst fractures (T11-L2) with or without suspected PLC injury. Pathological fractures, multilevel injuries, obvious posterior bony element injury and translation/dislocation injuries were excluded. Ten patients were randomly selected and commonly reported CT parameters indicating PLC injury, including superior inferior endplate angle, vertebral body height loss, local kyphotic deformity, interspinous distance and interpedicular distance were assessed for variability. Observer values were compared with an experienced gold rater in spinal trauma. Analysis of variability was performed for all observers, between the principal investigators and also between observers participating in each center.

Results: The studied parameters showed considerable variability in measurements among all observers and amongst all participating centers. The variability between the principal investigators was lower, but still substantial. The deviation of observer measurements from the gold rater were also significant for all CT parameters.

Conclusions: CT-based radiological parameters previously reported to be suggestive of PLC injury showed considerable variability and magnetic resonance imaging verification of a PLC injury in all doubtful cases is suggested.

Keywords

thoracolumbar fractures, spine trauma, posterior ligament complex, variability, reliability, computed tomography

[1] Ganga Hospital, Coimbatore, Tamil Nadu, India
[2] Assiut University Hospital, Assiut, Egypt
[3] The Uijeongbu St Mary's Hospital, Uijeongbu-si, South-Korea
[4] The Royal Victoria Hospital, Belfast, UK
[5] The Ondo State Trauma and Surgical Centre, Ondo State, Nigeria
[6] The First Hospital of Zhejiang University, Hangzhou, China
[7] Cajuru University Hospital, Curitiba, Brazil

Corresponding Author:
Shanmuganathan Rajasekaran, Department of Spine Surgery, Ganga Hospital, Mettupalayam Road, Coimbatore 641043, Tamil Nadu, India.
Email: rajasekarn.orth@gmail.com
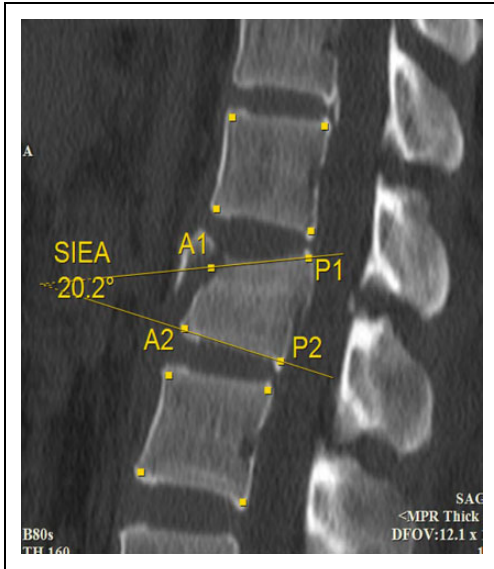
Figure 1. Measurement of superior inferior end plate angle (SIEA).

Figure 2. Measurement of anterior and posterior body heights (BH).

Figure 3. Measurement of local kyphotic angle (LK).

Figure 4. Measurement of interspinous distance (ISD).

Introduction

Thoracolumbar (TL) fractures are among the most common spine injuries encountered in emergency trauma care[1-3] with burst fractures accounting for 10% to 20% of all spinal fractures. These fractures usually result from high-energy trauma such as motor vehicle accidents or falling from heights.[4-9]

Treatment strategy for thoracolumbar fractures mainly depends on the stability of the fractured spine. Currently, there are various classification systems available to judge the stability of vertebral fractures and facilitate clinical decision making. While conservative treatments can be used for stable fracture, typically surgical treatments are suggested for unstable vertebral fractures.[10-12]

One of the most important stabilizing structures in the spine is the posterior ligament complex (PLC), which consists of the suprapinous ligament (SSL), interspinous ligament (ISL), ligamentum flavum (LF), and facet joint capsules. Flexion distraction injuries and Chance fractures typically lead to a failure of the anterior and middle column in compression and posterior column failure in distraction. These distraction forces lead to a disruption of the PLC resulting in an unstable spine injury.[10-12] PLC protects the spine against excessive flexion, extension, translation, and rotation. Injury to the PLC may lead to spinal instability, late spinal deformity, and persistent back pain and can jeopardize neural elements resulting in neurological deficit.[10,11] Assessing the integrity of the PLC has a bearing on the definitive management for these injuries.[13,14] Usually, PLC injury requires surgical fixation as it may result in progressive kyphosis of spine and poor functional outcome.[13-15]

Multidetector computed tomography scan (MDCT) is considered the method of choice to assess TL spine fractures and other visceral injuries in emergency trauma care.[16] However, PLC injury is difficult to assess on radiographs and CT scans, and ma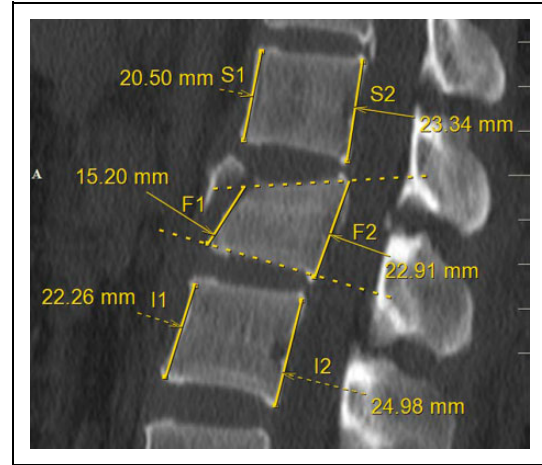gnetic resonance imaging (MRI) is considered to be the gold standard to identify PLC damage.[17-21] MRI assessment of PLC injuries has the disadvantages of a considerably longer scan duration, lower availability, poor feasibility in polytraumatized
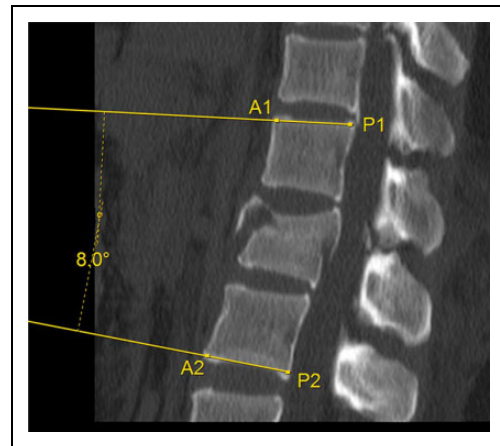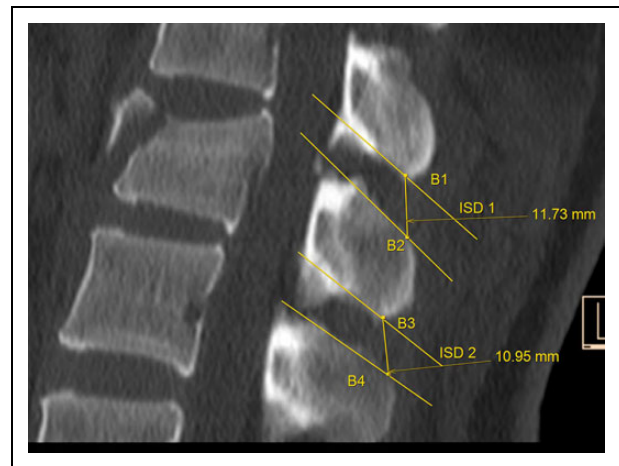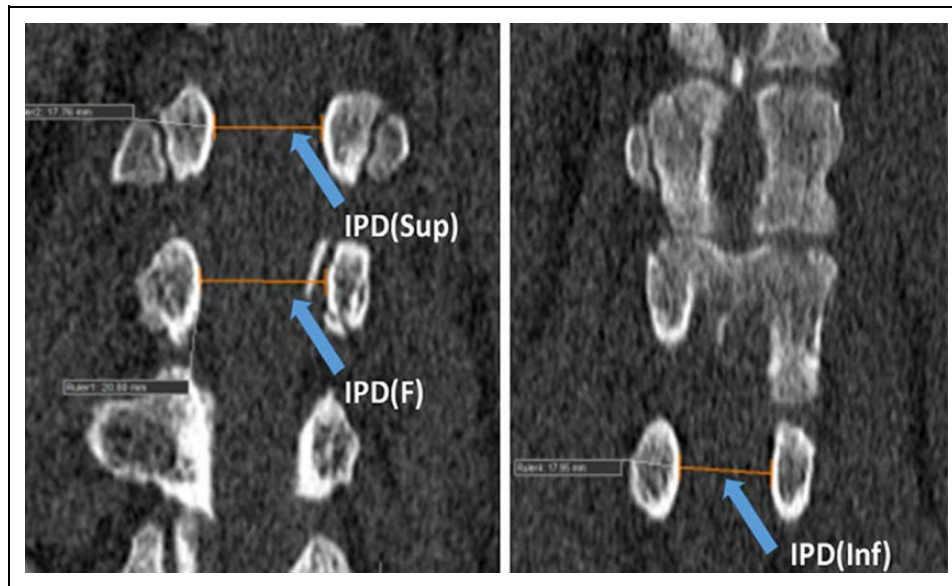
Figure 5. Measurement of interpedicular distance (IPD).

patients, and higher cost when compared with MDCT scans.[16-22] In contrast, CT scan is an integral part of trauma evaluation protocols, which is quick and accurate in the diagnosis of vertebral fractures diagnosis, particularly in the evaluation of polytrauma scenarios.[16,22,23] A number of authors have suggested to identify PLC injury on CT scans, based on criteria such as facet diastasis, local kyphosis, and interspinous distance.[12,24-28] The possibility to accurately identify PLC injury on CT scans would prove to be a great advantage, thus expediting treatment of TL fractures without MRI, especially in the clinical scenario of polytrauma with an unstable patient.[23]

Although reports have suggested that CT-based radiological parameters may reliably detect PLC injury.[15-20] To date, however, the variablility and reliability of these parameters have not been studied. Therefore, the aim of this study was to determine the interobserver variability and hence the feasibility of these CT based parameters to assess PLC injury compared with MRI when used in a multicenter setting.

Methods

A pilot study based on a retrospective multicenter case series was conducted at 7 tertiary referral centers (3 in Asia, 2 in Africa, 1 in Europe, and 1 in South America). The protocol was approved by the ethics committees or institutional review boards at each participating institution.

Each of the centers contributed multiple observers (N = 42) who evaluated the CTs. The observers' background was either spinal surgeon (N = 38) or radiologist (N = 4). In each of the centers, there was 1 principal investigator (PI) who attended 2 face-to-face training sessions on how to perform the measurements. These 2 training sessions were held by the gold rater at a time interval of 6 months apart. The gold rater was an experienced musculoskeletal radiologist (AM) with more than 10 years of experience in evaluating traumatic spine injury. The

training session was a day long workshop with the gold rater training the PIs on the image assessment protocol. The training sessions included ten patient image folders which were used for trial measurements for the PIs. The PIs were responsible to train their own site personnel. Once the training sessions were completed, an instructional video was sent to all sites, serving as a reference tutorial during the assessments, 6 weeks following the second face to face training session. Apart from the video, the investigation protocol document had a section dedicated to elaborating the steps in the measurement of the radiological parameters, which served as a written manual.

From the databases of each center, all adult patients (18-60 years old) with A3 or A4 TL burst fractures between T11 to L2 with or without suspected PLC injury who had undergone complete CT and MRI diagnostics before treatment and admitted to the hospital were identified and entered into our database (REDCap, Vanderbilt University, Nashville, TN, USA). Patients were excluded if they had (*a*) pathological fractures, (*b*) multilevel contiguous or noncontiguous injuries, (*c*) fractures with obvious spinous process split indicating tension band failure, or (*d*) fractures with translation injuries or dislocations which imply an obvious PLC injury.

Initially, sagittal and coronal reconstructed images of 91 patients obtained from a multislice CT scanner with volume acquisition by axial sections and a slice thickness of at least 1.25 mm were collected. Ten cases were randomly selected to evaluate various parameters that could potentially serve to determine PLC integrity as follows:

1. Superior inferior endplate angle (SIEA): The angle formed between the lines drawn along the superior and inferior endplate of the fractured vertebra (Figure 1).
2. Vertebral body height (BH): The vertebral body height was measured at 2 locations; first along the anterior vertebral body margin between the anterosuperior corner

Table 1. Parameter Measurement Protocol.

Parameter	Measurement Protocol
Landmarks	A1: corresponds to the anterior corner of superior endplate of the vertebral body A2: corresponds to the anterior corner of inferior endplate of the vertebral body P1: corresponds to the posterior corner of superior endplate of the vertebral body P2: corresponds to the posterior corner of inferior endplate of the vertebral body
Superior inferior endplate angle (SIEA) (Figure 1)	The measurement is performed on the mid-sagittal computed tomography (CT) reformatted image. A line is drawn along the superior (P1 to A1) and inferior endplates (P2 to A2) of the fractured vertebra. The angle formed between the lines is considered to be superior and inferior endplate angle.
Vertebral body height loss (BH) (Figure 2)	The measurement is performed on the mid-sagittal image of the CT scan. The measurement is performed at 2 locations, anterior vertebral body margin and posterior body margin. The anterior vertebral body height is measured between the anterosuperior corner (A1) and the anteroinferior corner (A2) of the fractured vertebral body. The posterior vertebral body height is measured at the posterior vertebral body margin between the posterosuperior corner (P1) and the posteroinferior corner (P2) of the fractured vertebral body. The measurements are repeated at the cephalad and caudal vertebrae. Ignoring the displaced anterosuperior and posterosuperior fragments, a line is drawn along the superior endplate of the largest fractured fragment to isolate A1 and A2.
	S1 and S2 indicate the anterior and the posterior vertebral margin length of the superior uninvolved vertebra. F1 and F2 indicate the anterior and the posterior vertebral margin length of the fractured vertebra. I1 and I2 indicate the anterior and the posterior vertebral margin length of the inferior uninvolved vertebra. The vertebral body height loss is calculated using the following formulae. Mean anterior vertebral body height $(MAVBH) = S1 + I1 / 2$ Mean posterior vertebral body height $(MPVBH) = S2 + I2 / 2$ Anterior vertebral body height loss $= MAVBH - F1$ Posterior vertebral body height Loss $= MPVBH - F2$
Local kyphosis (LK) (Figure 3)	The measurements is performed on the mid-sagittal image of the CT scan. The local deformity angle is defined as the angle subtended by a line drawn parallel to the superior endplate of the immediately superior vertebral body and that parallel to the inferior endplate of the immediately inferior vertebral body to the fractured vertebral body.
Interspinous distance (ISD) (Figure 4)	The measurement is performed on the mid-sagittal CT reformatted image. Line is drawn along the superior and inferior margins of the spinous process of the fractured vertebrae. Line is drawn along the inferior margin of the spinous process of the proximal uninvolved vertebra. Similarly, a line is drawn along the superior margin of the spinous process of the distal uninvolved vertebra. The most distal points on the spinous process in contact with the line are marked as B1, B2, B3, B4. The distance between B1 and B2 is measured and marked as ISD1. The distance between B3 and B4 is marked as ISD 2. The increase in the interspinous distance is calculated as the difference of ISD1 and ISD2
Interpedicular distance (IPD) (Figure 5)	The interpedicular distance is measured on the coronal section image. The distance between the medial borders of the pedicles is measured for the fractured vertebra, the caudal and the cephalad uninvolved vertebral bodies for comparison. Because of local kyphosis, if it is not possible to measure the interpedicular distance of all the 3 contiguous levels in a single section, it is measured in 2 consecutive sections.

and the anteroinferior corner of the fractured vertebral body and second along the posterior vertebral body margin between the posterosuperior corner and the posteroinferior corner of the fractured vertebral body. The loss in body height was assessed by comparing the fractured body height with the mean dimensions of the uninvolved superior and inferior vertebral body (Figure 2).

3. Local kyphotic deformity (LK): The angle formed between the lines drawn along the superior endplate of the cephalad and the inferior endplate of the caudal uninvolved vertebra (Figure 3).

4. Interspinous distance (ISD): The distance between the spinous process of the cephalad normal vertebra and the fractured vertebra measured and this distance was compared with the caudal uninvolved segment (Figure 4).

5. Interpedicular distance (IPD): This was measured as the distance between the medial borders of the 2 pedicles of the fractured vertebra. This was compared with the mean of the uninvolved adjacent segments cephalad and caudal to the fractured segment (Figure 5).

Further details on how the respective parameters were measured are presented in Table 1.

All CT measurements were made using Surgimap Spine (Nemaris, Inc, New York, NY, USA) a software program that has been validated for radiological parameter assessment.[29] The landmarks for measurement of radiological parameters were based on manual human measurement. Automatic contour recognition was not used as the fractured vertebra could not be accurately recreated with automated contour recognition. Analysis of variability was performed for all observers, between the 7 PIs and also between observers participating in each center.

Table 2. Descriptive Statistics of the Measurements Performed by 42 observers for the 5 Computed Tomography (CT) Parameters.

| | LK (Degree) | SIEA (Degree) | ISD (mm) | IPD (mm) | Percentage Loss of Body Height | |
					Anterior	Posterior
Case 01						
n	42	42	42	41[a]	42	42
Mean (SD)	4.24 (2.93)	14.17 (4.49)	1.44 (0.89)	2.14 (0.81)	41.04 (9.30)	19.55 (5.65)
Min; Max	0.93; 14.40	4.90; 22.77	0.00; 4.02	0.25; 3.71	6.48; 54.51	2.94; 32.01
Case 02						
n	42	42	42	42	42	42
Mean (SD)	16.43 (3.21)	11.28 (5.50)	3.42 (1.30)	2.17 (1.23)	38.63 (10.83)	25.05 (7.25)
Min; Max	2.00; 21.00	1.99; 28.00	0.64; 5.83	0.04; 5.00	−9.69; 57.92	4.52; 39.51
Case 03						
n	42	42	42	41[a]	42	42
Mean (SD)	27.39 (2.23)	32.99 (19.09)	1.80 (1.39)	5.64 (1.23)	51.80 (11.99)	−0.60 (8.39)
Min; Max	21.00; 31.50	13.87; 150.00	0.18; 6.74	0.70; 7.69	17.91; 77.76	−16.82; 22.61
Case 04						
n	42	42	42	42	42	42
Mean (SD)	10.10 (2.01)	10.29 (6.15)	1.65 (1.12)	1.67 (0.73)	28.78 (11.63)	19.19 (12.98)
Min; Max	5.00; 14.00	1.00; 28.65	0.01; 4.21	0.10; 3.25	−31.47; 49.51	−3.39; 51.29
Case 05						
n	40[a]	40[a]	39[a]	15[a]	40[a]	40[a]
Mean (SD)	15.40 (2.27)	13.26 (5.70)	1.36 (0.78)	2.79 (1.27)	32.44 (8.67)	15.00 (11.99)
Min; Max	8.30; 19.00	4.00; 23.00	0.10; 3.40	0.00; 4.95	−1.04; 47.39	−3.53; 43.92
Case 06						
n	42	42	42	42	42	42
Mean (SD)	18.64 (2.42)	16.63 (6.52)	3.09 (1.37)	2.03 (1.18)	39.07 (10.34)	15.29 (14.38)
Min; Max	9.00; 23.00	3.00; 24.14	0.19; 6.43	0.21; 8.07	7.99; 59.17	−6.34; 43.56
Case 07						
n	42	42	42	42	42	42
Mean (SD)	2.18 (2.36)	12.76 (6.98)	1.31 (1.61)	0.73 (0.80)	38.45 (15.62)	9.83 (13.27)
Min; Max	0.00; 14.00	1.00; 31.90	0.03; 9.65	0.00; 4.89	−15.03; 79.71	−10.37; 54.49
Case 08						
n	42	42	42	41[a]	42	42
Mean (SD)	15.28 (2.13)	20.81 (5.77)	2.91 (2.58)	7.79 (1.44)	58.82 (13.76)	9.46 (11.52)
Min; Max	8.00; 19.10	7.33; 31.00	0.00; 13.82	4.06; 11.44	−13.59; 74.70	−16.03; 35.26
Case 09						
n	42	42	42	41[a]	42	42
Mean (SD)	26.89 (2.14)	29.77 (4.26)	15.03 (2.58)	1.89 (1.89)	51.25 (13.51)	-10.15 (11.36)
Min; Max	21.00; 31.00	16.00; 39.00	9.07; 20.19	0.00; 8.85	−8.25; 71.39	−28.47; 24.57
Case 10						
n	42	42	42	42	42	42
Mean (SD)	8.44 (2.76)	22.52 (7.34)	4.05 (2.46)	6.37 (1.28)	55.97 (15.32)	17.22 (13.24)
Min; Max	0.80; 14.10	3.00; 36.00	0.89; 15.66	1.39; 8.80	−7.94; 71.66	−15.32; 52.72

Abbreviations: CT, computed tomography; LK, local kyphosis; ISD, interspinous distance; IPD, interpedicular distance.
[a] Observer was not able to assess measurements for all cases.

Statistical Methods

We conducted several analyses to assess the variability between the observers, between the PIs and between the observers at the different participating centers. Under the assumption that the gold rater's measurements would be the most accurate of all, we also compared the measurements between the gold rater and the remaining observers. This was done by determining the mean absolute deviation (MAD) from the gold rater. The MAD is the sum of the absolute difference between the measurement of each individual observer and the measurement of the gold rater, divided by the number of observers. Analysis of variability was performed for all

observers, between the PIs and also between observers participating in each center using presentations of mean, standard deviation, minimum and maximum values and graphical by using box plots. Intraclass correlation was tested with Shrout-Fleiss reliability coefficient for all observer and for PI separately.[30]

Results

There were 42 observers who performed measurements on the CT scans of 10 cases. In most cases, all observers were able to analyse all parameters for each case. However, in a few cases,

Table 3. Mean Absolute Deviation (MAD) Assessed by All Observers (n = 41) From the Gold Rater for the 5 CT Parameters.

| | | | | | Percentage Loss of Body Height | |
| | | | | | MAD From the Gold Rater[a] | |
	LK (degree)	SIEA (degree)	ISD (mm)	IPD (mm)	Anterior	Posterior
Case 01	2.58	5.12	1.00	0.70	6.86	12.10
Case 02	3.18	5.17	1.08	2.04	9.33	6.60
Case 03	1.77	6.46	1.57	—[a]	9.50	8.07
Case 04	1.84	5.18	1.57	0.96	6.81	19.21
Case 05	2.88	7.68	0.81	—[a]	6.06	10.58
Case 06	1.84	7.56	1.10	1.50	10.43	14.79
Case 07	1.76	12.05	1.14	0.99	10.97	9.56
Case 08	2.77	7.13	2.07	—[a]	8.72	8.64
Case 09	4.21	3.34	2.16	1.13	15.59	9.26
Case 10	3.06	7.51	1.43	1.03	16.07	13.36
Overall	2.59	6.72	1.40	1.19	10.05	11.22

Abbreviations: CT, computed tomography; LK, local kyphosis; ISD, interspinous distance; IPD, interpedicular distance.
[a] Gold rater was not able to assess IPD measurements for Cases 03, 05, and 08.

observers and gold rater were unable to perform the measurements either due to difficulties in assessing the landmarks outlined in the measurement protocol or due to technical issues where many raters were unable to calibrate the images in Surgimap, for example, case 5, which has been marked with a superscript "a" in Tables 2 and 3. Table 2 shows the descriptive statistics for the measurements performed by 42 observers for the 5 CT parameters.

Throughout all CT parameters, we found a very high variability in the individual measurements. The variability was most pronounced when comparing the measurements among all observers as well as amongst all teams. The variability among the PIs was lower than among all observers, but still considerable (Figures 6–8).

The deviation of all observers measurements from the gold rater's measurements was also substantial for all 5 CT parameters (Figure 9). The average MAD from the gold rater ranged from 1.19 for IPD to 11.22 for posterior BH. The average MAD for kyphosis measurements was 2.59 for LK and 6.72 for SIEA (Table 3).

The interclass correlation coefficients generated by the Shrout-Fleiss reliability coefficient analysis are reported for the selected CT parameters in Table 4. LK, IPD, and ISD showed good reliability with values between .75 and .90 for both the PI group and other observers.

Discussion

The management principles in TL fractures are based on the assessment of associated neurological injury, potential spinal instability and possibility of development of late-onset spinal deformity.[1,3,10,31] Management guidelines of TL fractures proposed by both the Spine Trauma Study Group (TLICS classification) and AOSpine proposed thoracolumbar spine fracture classification is influenced by the presence/absence of the PLC injury.[15,32] The PLC includes the supraspinous ligament (SSL),

interspinous ligament (ISL), ligamentum flavum (LF), and facet joint capsules, and these structures are best assessed on MRI scans.[24,33] The ligamentous healing potential is considered poor; therefore, an accurate identification of PLC injury is necessary to allow optimal management of TL fractures.[10,11]

TL fractures can be associated with other more life threatening visceral organ injuries in the thorax and abdomen.[20] These injuries can render the patient hemodynamically unstable, thus making an early MRI for assessment of PLC challenging if not impossible. CT scans score over MRI as the preferred modality for evaluation of spinal trauma due to a quick turnover time and can be performed even in an unstable clinical scenario.[16,22,28] The CT images can provide clear details to identify fracture morphology, spinal canal compromise, and unstable spine fractures. The possibility to identify PLC injuries on CT scans would prove of great clinical benefit, as it would likely reduce the evaluation time and cost and thus expedite treatment in TL spine fractures. Based on previously reported radiological criteria, our study attempted to assess the interrater variability of measurements based on CT scans.

Various authors have proposed radiological criteria to assess PLC injury based on radiographs and CT scans.[24,27] The drawback of these reports is that they are all single center studies and information on the reliability of the findings in a multicenter setting is lacking. Hiyama et al[24] assessed loss of vertebral body height, local kyphosis, vertebral body translation, canal compromise, interlaminar distance, supraspinous distance, and interspinous distance in 40 patients with TL fractures and concluded that a local kyphosis >20° and increased supraspinous distance were associated with PLC injury. Rajasekaran et al[25] analyzed 60 patients with possible PLC injury using STIR MRI sequences as a gold standard to identify PLC injury. The authors reported that local kyphosis of >20° and interspinous distance increase of ≥2 mm compared with adjacent levels may serve as a radiological criterion to predict PLC injury, especially in the scenario of emergency trauma where
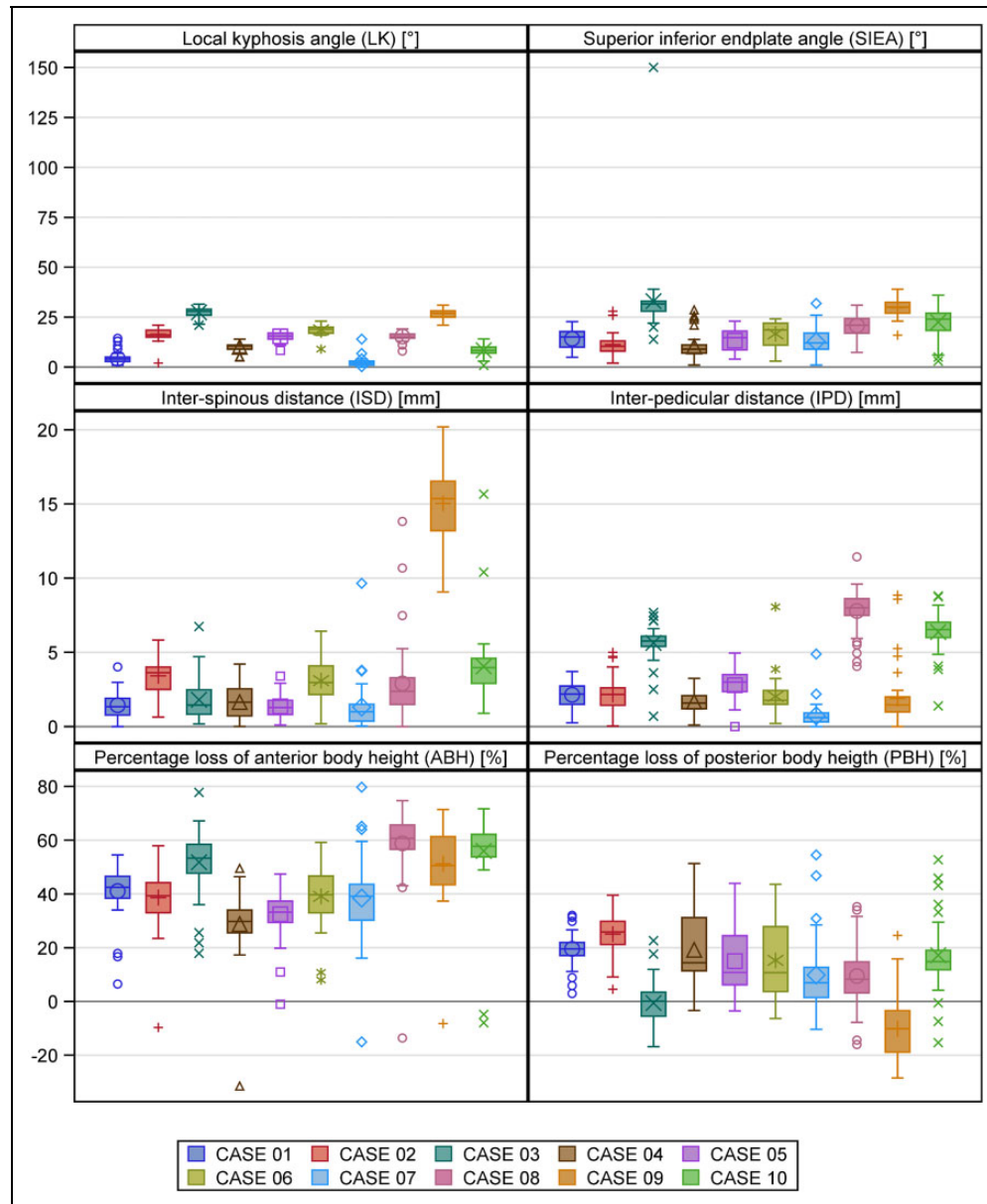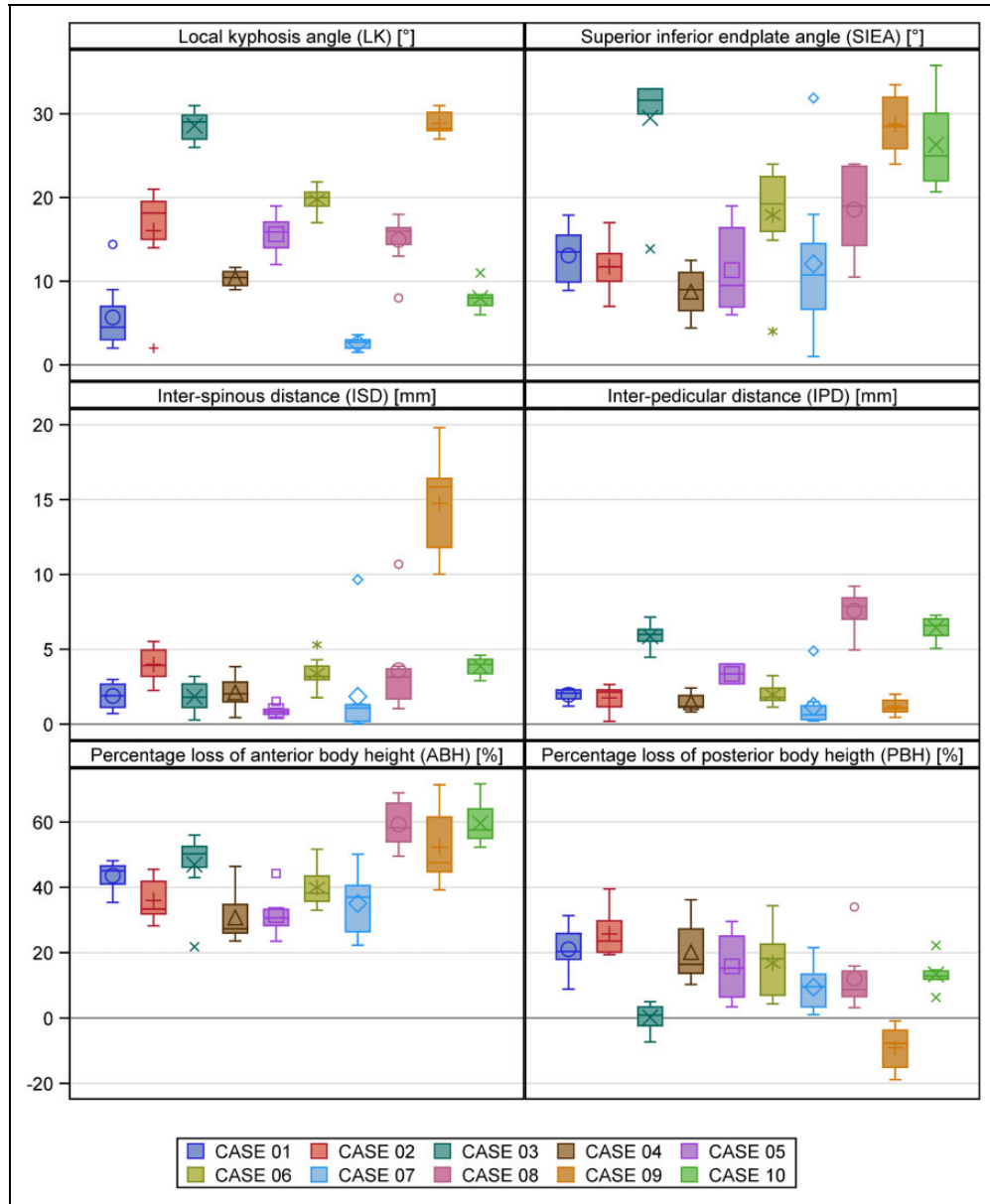
Figure 6. Box plot for the computed tomography parameters per case: all observers (n = 42).

MRI is not feasible.[25] Barcelos et al[27] noted that measurements from CT scans could reliably predict PLC injury and an increase in the interspinous distance could be used to differentiate AO type A versus type B injuries. Yee et al[28] reported on a survey conducted among members of the spine trauma study group for radiological predictors of PLC injury. The survey concluded that obvious translation as seen in AO type C injuries and interspinous distance ≥7 mm was felt to be the most reliable indicator of PLC disruption. The members felt that diastasis of the facet on CT scans was the best indicator of PLC injury, especially when the plain radiographs appeared normal. However, there was no consensus on any particular radiological parameters to suggest PLC injury.[28]

Our study was based on measurements of 42 observers from 7 different centers, which is a considerably higher number of

observers than in previous reports on the predictive value of CT-based parameters for the assessment of PLC damage. Therefore, it offers a more robust assessment of the reliability of CT scan images to predict PLC injury. Our study also included both radiologists and spine surgeons as observers, which is closer to a real-life setting.

We saw a considerable variability in the reported measurements for all radiological parameters. The variability was present amongst all observers as well as among the different centers. Of note, the variability among the PIs alone was slightly lower, suggesting suboptimal transfer of knowledge to the remaining observers. However, the variability even amongst the PIs was high enough to preclude using the proposed CT-based parameters to accurately assess PLC injury.

Figure 7. Box plot for the computed tomography parameters per case: all principal investigators (n = 8).

A possible explanation for the significant variability was the difficulty encountered in identifying the landmarks for the measurements. In particular, identification of the anterior and posterior vertebral margin was challenging due to the presence of retropulsion at the posterosuperior corner and displacement of fragments at the anterosuperior corner.

The effect of the difficultt to accurately identify the landmarks on the endplates is also illustrated by the different degrees of variability seen with the measurement of SIEA and LK. Both parameters are measures for kyphotic deformity. LK, which was measured between the superior and inferior uninvolved vertebral body, showed less variability than SIEA, which was measured directly on the fractured vertebra. This clearly shows that measurements taken from intact endplates, which require less interpretation from the observer's side, offer

less potential for different interpretation than measurements taken from endplates with retropulsed or otherwise displaced fragments. This probably resulted in the greater uniformity of results for LK compared to SIEA. The authors feel that, although findings such as increased interspinous distance and local kyphosis have been reported to be indicative of PLC disruption, an objective definition to quantify the change in radiological measurement is difficult to develop. There is considerable natural variation in the individual fracture patterns and spinous process anatomy, which may preclude development of an objective cutoff or measurement protocol and eventual prediction of PLC injury on CT scans.

The absolute values for parameters such as ISD and IPD were very small, so that the MAD from the observers to the gold rater ranged between 1 and 3 mm. With such small values,
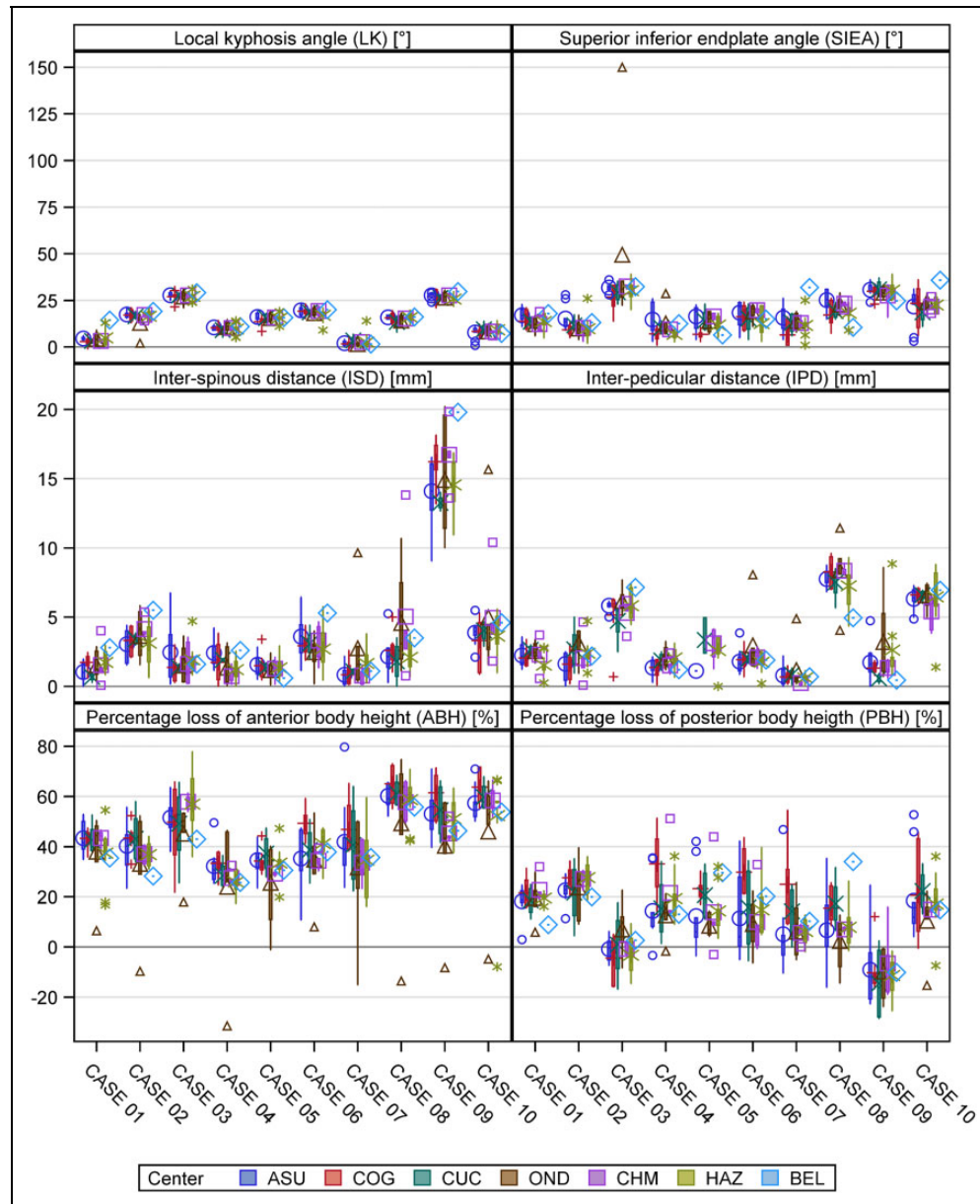
Figure 8. Box plot for the computed tomography parameters per case per site.

minor deviations on setting the landmark points may lead to an extremely high variability of measurements especially as the measurements for the landmark were performed manually. Additionally, for one case the images could not be calibrated by the majority of observers, including the gold rater. In an attempt to ensure uniform measurements, every observer had received detailed instructions on how to identify these landmarks and personal hands-on training had been provided to the PIs of each site. However, the considerable variability of results suggests that either the instructions were insufficient, the Surgimap software may have been unsuitable to perform the measurements with sufficient accuracy, or the proposed predictive parameters could be unsuitable for the assessment of PLC integrity altogether.

Limitations

An important limitation of the study is that the reference measurements were not objective, because they were also derived from personal judgement. Therefore, human error could have also been present in the reference measurement. Another difficulty was that the experience of the observers with regard to CT evaluation was very heterogeneous. On one hand, this may have contributed to the low reproducibility of results. On the other hand, this reflects exactly how evaluation of the proposed radiological parameters is done in clinical practice, which increases the generalizability of the results we found. There were 42 observers however, the radiologist to surgeon ratio was 4:38, which could have been more comparable. The interclass correlation coefficients for LK, IPD and ISD showed
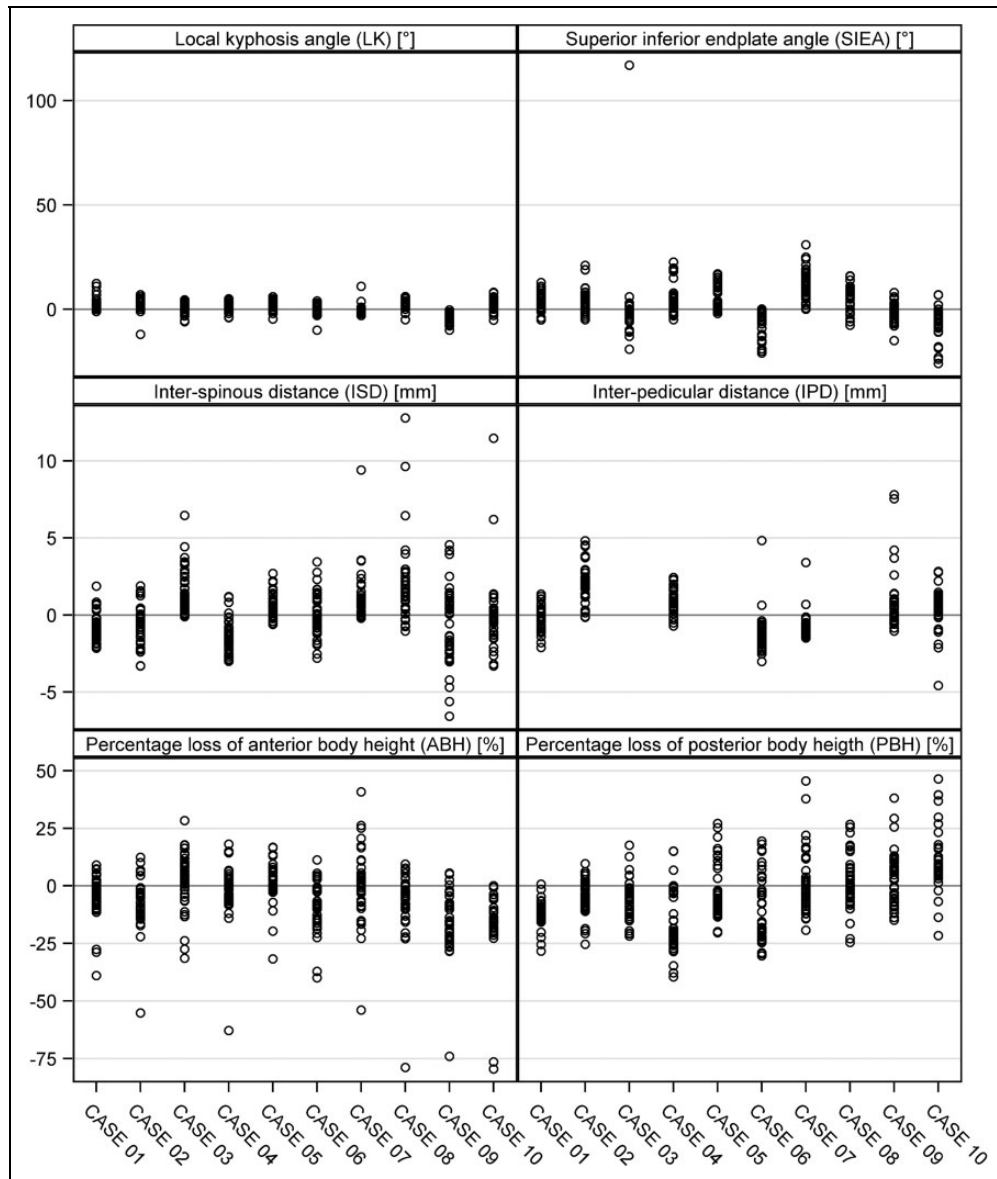
Figure 9. Deviation from gold rater per case.

Table 4. Intraclass Correlation Coefficients of CT Measurements

Parameter	No. of Cases	Shrout-Fleiss Reliability Coefficient for Observers (n = 42)	Shrout-Fleiss Reliability Coefficient for Principal Investigators (n = 8)
LK	10	.930	.916
SIEA	10	.488	.680
ISD	10	.852	.809
IPD	10	.786	.867
BH (anterior)	10	.591	.704
BH (posterior)	10	.523	.667

Abbreviations: CT, computed tomography; LK, local kyphosis; ISD, interspinous distance; IPD, interpedicular distance; BH, body height.

good reliability. However, since the number of cases is small the statistical power of the analysis is low and the authors would advice caution in the interpretation of these values.

Conclusions

The CT parameters local kyphosis (LK), superior inferior end plate angle (SIEA), vertebral body height (BH), interspinous distance (ISD) and interpedicular distance (IPD) have been proposed to be used to determine PLC integrity without the need for MRI. When assessing the inter-rater variability of these parameters in a multicenter setting with 42 observers of heterogeneous experience from 7 different centers from 4 continents, a high variability of results was seen. Therefore,

previous reports could not be validated. The study did not allow to determine whether the proposed parameters are unsuitable for PLC evaluation per se or whether insufficient knowledge transfer and an unsuitable measurement method were responsible for the poor reproducibility. Modified instructions to identify the landmarks as well as face to face training of each observer may help improve measurement uniformity. We conclude that currently, MRI verification of a PLC injury should be done in all doubtful cases.

ORCID iD

Osama Farouk https://orcid.org/0000-0002-3897-6485
Mohammad El-Sharkawi https://orcid.org/0000-0001-6177-7145

References

1. Wood KB, Li W, Lebl DR, Lebl DS, Ploumis A. Management of thoracolumbar spine fractures. *Spine J*. 2014;14:145-164.
2. el-Khoury GY, Whitten CG. Trauma to the upper thoracic spine: anatomy, biomechanics, and unique imaging features. *AJR Am J Roentgenol*. 1993;160:95-102.
3. Rajasekaran S, Kanna RM, Shetty AP. Management of thoracolumbar spine trauma: an overview. *Indian J Orthop*. 2015;49: 72-82.
4. 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.
5. Esses SI, Botsford DJ, Kostuik JP. Evaluation of surgical treatment for burst fractures. *Spine (Phila Pa 1976)*. 1990;15:667-673.
6. Gertzbein SD. Scoliosis Research Society. Multicenter spine fracture study. *Spine (Phila Pa 1976)*. 1992;17:528-540.
7. 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.
8. Kraemer WJ, Schemitsch EH, Lever J, McBroom RJ, McKee MD, Waddell JP. Functional outcome of thoracolumbar burst fractures without neurological deficit. *J Orthop Trauma*. 1996; 10:541-544.
9. Müller U, Berlemann U, Sledge J, Schwarzenbach O. Treatment of thoracolumbar burst fractures without neurologic deficit by indirect reduction and posterior instrumentation: bisegmental stabilization with monosegmental fusion. *Eur Spine J*. 1999;8: 284-289.
10. Kepler CK, Felte RF, Rihn JA. Current concepts: classification of thoracolumbar fractures. *Semin Spine Surg*. 2012;24:210-215.
11. Smith HE, Anderson DG, Vaccaro AR, et al. Anatomy, biomechanics, and classification of thoracolumbar injuries. *Semin Spine Surg*. 2010;22:2-7.
12. Vaccaro AR, Lee JY, Schweitzer KM Jr, et al; Spine Trauma Study Group. Assessment of injury to the posterior ligamentous complex in thoracolumbar spine trauma. *Spine J*. 2006;6:524-528.
13. Holdsworth F. Fractures, dislocations, and fracture-dislocations of the spine. *J Bone Joint Surg Am*. 1970;52:1534-1551.
14. Oner FC, van Gils APG, Faber JA, Dhert WJ, Verbout AJ. Some complications of common treatment schemes of thoracolumbar spine fractures can be predicted with magnetic resonance imaging: prospective study of 53 patients with 71 fractures. *Spine (Phila Pa 1976)*. 2002;27:629-636.
15. 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;30:2325-2333.
16. Brown CVR, Antevil JL, Sise MJ, Sack DI. Spiral computed tomography for the diagnosis of cervical, thoracic, and lumbar spine fractures: its time has come. *J Trauma*. 2005;58:890-896.
17. 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.
18. Haba H, Taneichi H, Kotani Y, 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.
19. van Middendorp JJ, Patel AA, Schuetz M, Joaquim AF. The precision, accuracy and validity of detecting posterior ligamentous complex injuries of the thoracic and lumbar spine: a critical appraisal of the literature. *Eur Spine J*. 2013;22:461-474.
20. 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;12:391-396.
21. Morais DF, de Melo Neto JS, Meguins LC, Mussi SE, Filho JRLF, Tognola WA. Clinical applicability of magnetic resonance imaging in acute spinal cord trauma. *Eur Spine J*. 2014;23: 1457-1463.
22. Wurmb TE, Frühwald P, Hopfner W, et al. Whole-body multislice computed tomography as the first line diagnostic tool in patients with multiple injuries: the focus on time. *J Trauma*. 2009;66: 658-665.
23. Hessmann MH, Hofmann A, Kreitner KF, Lott C, Rommens PM. The benefit of multislice CT in the emergency room management of polytraumatized patients. *Acta Chir Belg*. 2006;106:500-507.

24. Hiyama A, Watanabe M, Katoh H, Sato M, Nagai T, Mochida J. Relationships between posterior ligamentous complex injury and radiographic parameters in patients with thoracolumbar burst fractures. *Injury*. 2015;46:392-398.

25. Rajasekaran S, Maheswaran A, Aiyer SN, Kanna R, Dumpa SR, Shetty AP. Prediction of posterior ligamentous complex injury in thoracolumbar fractures using non-MRI imaging techniques. *Int Orthop*. 2016;40:1075-1081.

26. Chen JX, Goswami A, Xu DL, et al. The radiologic assessment of posterior ligamentous complex injury in patients with thoracolumbar fracture. *Eur Spine J*. 2017;26:1454-1462.

27. Barcelos ACES, Joaquim AF, Botelho RV. Reliability of the evaluation of posterior ligamentous complex injury in thoracolumbar spine trauma with the use of computed tomography scan. *Eur Spine J*. 2016;25:1135-1143.

28. Lee JY, Vaccaro AR, Schweitzer KM Jr, et al. Assessment of injury to the thoracolumbar posterior ligamentous complex in the setting of normal-appearing plain radiography. *Spine J*. 2007;7:422-427.

29. Akbar M, Terran J, Ames CP, Lafage V, Schwab F. Use of Surgimap Spine in sagittal plane analysis, osteotomy planning, and correction calculation. *Neurosurg Clin N Am*. 2013;24:163-172.

30. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-163.

31. Dai LY, Jiang SD, Wang XY, Jiang LS. A review of the management of thoracolumbar burst fractures. *Surg Neurol*. 2007;67: 221-231.

32. Vaccaro AR, Oner C, Kepler CK, et al; AOSpine Spinal Cord Injury & Trauma Knowledge Forum. AOSpine thoracolumbar spine injury classification system: fracture description, neurological status, and key modifiers. *Spine (Phila Pa 1976)*. 2013;38: 2028-2037.

33. 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.