How Often Would MRI Change the Thoracolumbar Fracture Classification or Decision-Making Compared to CT Alone?

Mohamed M. Aly, MD, MSc, PhD[1,2] ⬤, Abdulbaset M. Al-Shoaibi, MD, MSc, FRCR, MD[3],
Saleh Abduraba Ali, MD[3] ⬤, Areej Al Fattani, MPH, CCRP[4] ⬤, and
Hany Eldawoody, MD, PhD[1,2]

Abstract

Study design: retrospective study of consecutive patients

Objective: to analyze the frequency of change in Thoracolumbar fractures (TLFs) classification or decision-making after MRI compared by CT alone.

Methods: A retrospective review of 244 consecutive patients with acute TLFs (T1-L5) presented to a single level 1 trauma center between 2014 and 2021. Three and 4 reviewers independently classified all fractures according to AOSpine and AOSpine injury severity score (TLAOSIS) by CT then MRI, respectively. Posterior ligamentous complex Injury (PLC) was diagnosed on CT and MRI by ≥ 2 positive CT findings and Black stripe discontinuity.

Results: MRI changed AO classification in 25/244 patients (10.2%, $P < .0001$) due to an 8.2% upgrade from type A to type B and a 2% downgrade from type B to type A. The addition of MRI changed TL AOSIS among the 3 treatment recommendation groups in 35/244 (19.7%, 95% CI [14.9%-25.2%]. The best predictor of upgrade from type A to type B and downgrade from type B to type A was a single positive CT finding and the presence of only 2 CT signs as opposed to ≥3 signs, respectively ($P < .0001$ $P = .03$, respectively). Thoracic fractures showed a significantly higher reclassification rate than thoracolumbar and low lumbar (20% vs 10% and 0%, respectively, $P = .07$).

Conclusion: using appropriate CT/MRI criteria for PLC injury, MRI changed the AOSpine classification by 10% and TLAOSIS based treatment by 19.7%. The best predictors of fracture reclassification by MRI were the number of positive CT findings and fracture level.

Keywords

computed tomography, thoracolumbar fractures, magnetic resonance imaging, AOSpine classification, thoracolumbar injury classification

Introduction

Over the last few decades, several classification systems have proposed classifying thoracolumbar fractures (TLFs) based on injury's morphology or mechanism as depicted on plain radiography or computed tomography (CT).[1,2] More recently, the Thoracolumbar Injury Classification System (TLICS) and the new AOSpine classification systems have emphasized the integrity of the posterior ligamentous complex (PLC).[3,4] As a result, the use of magnetic resonance imaging (MRI) as a reference study for PLC

[1] Department of Neurosurgery, Prince Mohammed Bin Abdulaziz Hospital, Riyadh, Saudi Arabia
[2] Department of Neurosurgery, Mansoura University, Mansoura, Egypt
[3] Department of Diagnostic Radiology, Prince Mohammed Bin Abdulaziz Hospital, Riyadh, Saudi Arabia
[4] Department of Biostatistics Epidemiology and Scientific computing, King Faisal Specialist Hospital and Research Hospital, Riyadh, Saudi Arabia

Corresponding Author:
Mohamed M. Aly, Department of Neurosurgery, Prince Mohamed Ben Abdulaziz Hospital, P.O Box 54146, Riyadh 11514, Saudi Arabia.
Email: drmoali26@gmail.com and ali@pmah.med.sa

evaluation is growing.[5] However, its high cost, limited accessibility, and prolonged scanning time limit MRI's applicability in emergency settings and may explain why the new AOSpine classification was predominantly CT-based.[4,5] On the other hand, the previously reported high rate of PLC injury overlooked on CT alone remains a major concern.[6] There are no clear guidelines for MRI in neurologically intact patients with TLFs, resulting in missed PLC injuries or, at times, overuses of MRI.[7]

According to a recent systematic review, only 2 previous studies have indicated that MRI can significantly alter TLFs' classification or management plan (range 24-31% and 16%, respectively).[7-9] However, previous research might have had significant methodological bias and provided insufficient evidence to guide decisions about when to perform MRI for TLFs.[7] First, the small sample size (n = 33 and 100) and non-consecutive patient recruitment could have produced a sampling bias, limiting the findings' generalizability.[7-10] More importantly, the previous work could have resulted in a misclassification bias due to imperfect MRI or CT criteria for identifying PLC injury.[10,11] Previous studies defined PLC injury on MRI as any abnormal MRI signal change, rather than black stripe discontinuity, even though this criterion has been shown to "over-read" PLC injuries compared to intraoperative findings.[12,13] Similarly, previous studies defined PLC injury on CT based on any positive CT findings of PLC injury without accounting for the varied predictive values or combined value of CT findings for PLC injury.[8,9,14] Notably, 2 recent studies examined the diagnostic value of combined CT findings concluding that at least 2 positive CT findings, but not a single finding, yielded a sufficiently high predictive value (88-91%) to warrant its use as a criterion for PLC injury.[15,16] Finally, the Thoracolumbar AOSpine Injury Severity Score (TLAOSIS) has proposed treatment recommendations based on injury score: ≤ 3 conservatives, 4-5 either surgical or conservative, > 5 surgical. However, no prior studies have evaluated the impact of MRI in decision-making based on TLAOSIS compared with CT alone.[17] This study's primary goal is to examine the impact of MRI on TLFs' classification and decision-making compared to CT alone, attempting to address most of the shortcomings of previous studies.

Material and Methods

Ethical Statement

Our institutional review board approved this single-center study, which was conducted in a level 1 trauma center. Informed consent was waived due to the retrospective design of the study. We have followed the published guidelines on the Standards for the Reporting of Diagnostic Accuracy Studies (STARD).[18,19]

Inclusion/Exclusion Criteria

We retrospectively reviewed our radiology information system for consecutive patients who presented to our center with 1 or more acute TLFs between January 2014 and June 2021. All patients suspected of high-energy spinal trauma in our center perform a whole-body CT. Additionally, most of the patients will do MRI, but the decision may vary according to the preference of the primary physician. The inclusion criteria were (1) a thoracic or lumbar vertebral body fracture (T1–L5), excluding patients with isolated posterior element fractures; (2) patients who underwent both CT and MRI within 10 days of the date of the Injury[20] 3) patients with multi-level fractures were included; however, we only graded the injured level with the highest AO grade.[4] Of 365 patients who fulfilled the inclusion criteria, we excluded 121 patients (33.1%, Figure 1) for any of the followings: patients with a neurological deficit including complete or incomplete spinal cord or cauda equina or radicular symptoms (n = 84; including 43 type C fractures, 22 type B fractures,18 A4, and 1 A3 fracture; all were treated surgically),[5] osteoporotic fractures (n = 24; considered for low-energy trauma), inadequate coverage of fractured segments by MRI or CT images (n = 3), low-quality MRI images due to motion artifacts (n = 3), cervicothoracic fractures (n = 4), questionable fracture acuity (n = 1, suggested by lack of bone marrow edema on MRI), pathological fractures due to spinal metastases (n = 1), and prior spinal surgery (n = 1).

Data Collection

We collected the demographic data from the patients' records and included age, sex, mechanism of injury, levels and number of fractures, associated injuries, treatment mode, and neurological status (based on the American Spinal Injury Association [ASIA] Impairment Scale).[4] In addition, we conducted a longitudinal chart review for patients with MRI reclassification to determine whether they were treated surgically or conservatively. Cases treated conservatively were examined at the latest follow-up to see if they had back pain, deformity, neurological deficit, or required additional surgery.

CT and MRI Image Acquisition

Per previously described CT protocols, CT images were obtained using a 64-slice multidetector CT scanner (GE, Discovery HD 750, Chicago, Illinois, USA).[16] MRIs were performed on 2 different scanners: 125 patients on a 1.5-T MRI scanner (GE Optima MR450W) and 119 patients on a 3-T MRI scanner (Magnetom Skyra [Siemens Healthcare, Munich, Germany]). The MRI spine trauma protocol included axial and sagittal T2-weighted and T1-weighted images and sagittal short tau inversion recovery (STIR), with a slice thickness of 3 mm for sagittal and 4 mm axial, using a matrix size of 240 × 320.[21,14]

Image Interpretation Protocol

Three reviewers reviewed all CT and MRI images on PACS: a fellowship-trained neuroradiologist (AA) and 2 senior spinal surgeons (MMA, HE). A fourth reviewer, a general radiologist (SA), only looked at the MRI images. All 4 reviewers had more than 10 years of trauma imaging experience and prior consensus training on the standardization of the definition and
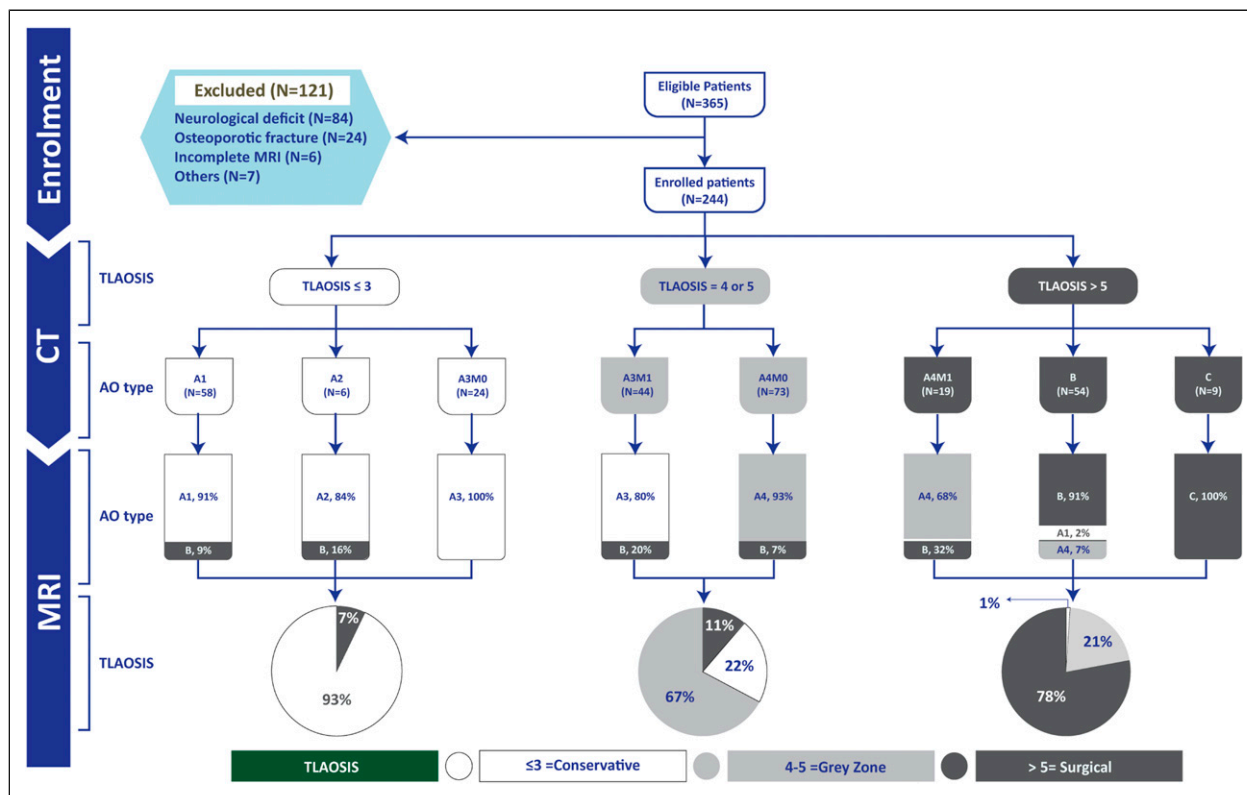
Figure 1. Flow chart of all participants in the study. *M1*, Modifier indicates indeterminate posterior ligamentous complex injury as defined by a single positive CT finding; *M0*, No positive CT findings and *TLAOSIS*, Thoracolumbar AOSpine injury score.

best imaging plan or sequence for each CT or MRI imaging fining (Figures 2 and 3).[16] Each reviewer independently classified fractures according to AOSpine Classification on CT and MRI in 2 sessions with 8 week intervals.[18,19] To calculate intra-rater agreement, a reviewer reviewed the images for 2 rounds. Clinical data and other readings were blinded to the reviewers during the image review.[3,18,19] We used the approval by a minimum of 2 of the 3 PLC status reviewers in CT and 3 of the 4 PLC status reviewers in MRI as the reference standard for diagnostic accuracy analysis. Disagreement about the status of the MRI (2 disrupted vs two non-disrupted) was resolved by a consensus decision among the 4 reviewers (default analysis).[22]

CT and MRI Grading of PLC Injury

We looked for the following 4 CT signs independently associated with PLC injury:[16,15,23] (1) facet joint malalignment (FJM), (2) horizontal fracture of the lamina or pedicle (HLF), (4) avulsion or transverse fracture of the spinous process (SPF), and (5) interspinous distance widening (ISW) >4 mm. (Figure 4E-4H). The number of positive findings on CT was used to classify PLC status as previously described: (1) Disrupted PLC if ≥ 2 positive CT findings, (2) Suspected PLC injury (M1 modifier) if 1 positive CT finding, (3) intact (M0) no positive CT findings.[15,16] We graded PLC integrity on MRI

as described by Pizones et al: (1) complete disruption, defined as Supraspinous ligament (SSL) or Ligamentum flavum (LF) disruption identified by a discontinuous black stripe; (2) non-disrupted, no MRI signal change or high signal intensity (HSI) due to interspinous ligament (ISL) edema or facet joint effusion (Figure 4A-4D).[16,24]

CT and MRI AOSpine Classification and the Thoracolumbar AOSpine Injury Score (TLAOSIS)

We classified all fractures in CT according to AOSpine Classification as follows; (1) type A (compression fracture with no evidence of translation or PLC disruption [intact or suspected PLC injury, ≤1 positive CT finding]; subtypes A1-A4); (2) type B (vertebral fracture with disrupted PLC [≥ 2 positive CT finding]); and (3) Type C (vertebral fracture with evidence of translation/rotation).[4] For type-A fractures, M1 modifier (suspected PLC injury) was defined if a single positive CT finding and M0 if no positive CT findings(Figure 5)[4,16]

We classified all fractures in MRI according to AOSpine Classification as follows: A compression fracture with complete PLC disruption was classified as type B, whereas that associated with non-disrupted PLC (intact PLC or HSI) was diagnosed as type A (no M1 modifier was assigned after MRI, Figure 5).In addition, CT and MRI were used to generate
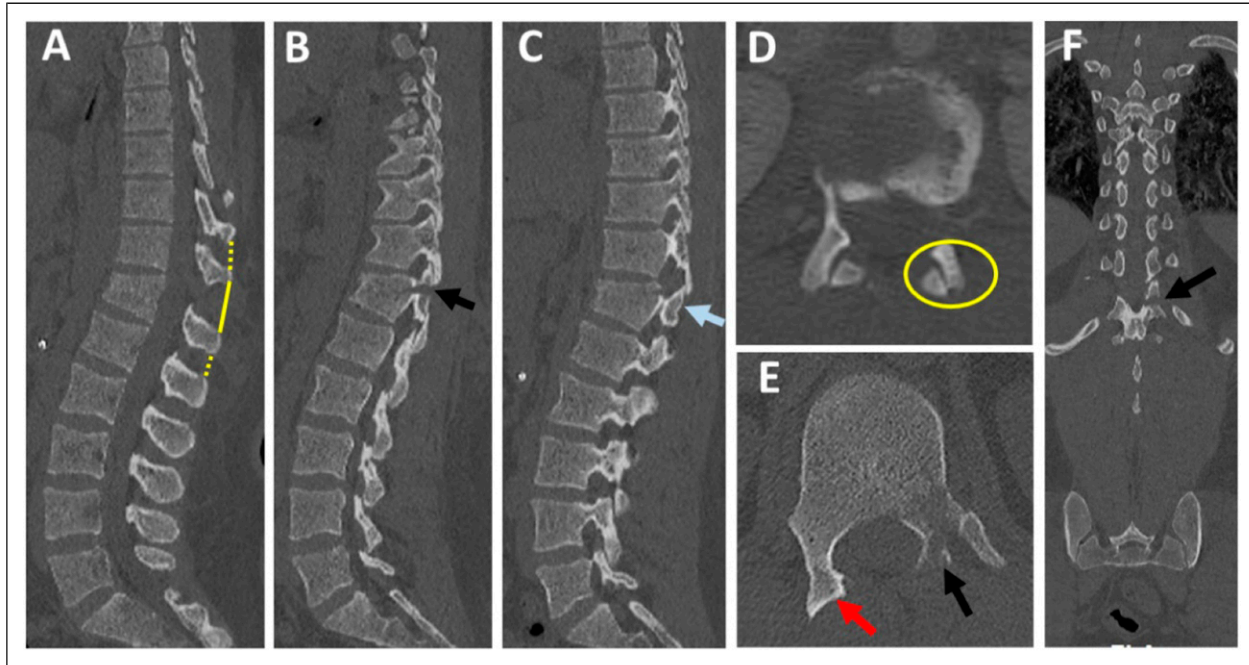
Figure 2. CT checklist for diagnosis of PLC injury. **(A)** Midsagittal CT shows should be checked for interspinous widening (yellow solid and interrupted lines, as well as vertebral translation and spinous process fracture; **(B-C)** parasagittal CT, shows left laminar and pedicle fracture (**B**, black arrow) and right subluxed facet (**C**, blue arrow); **(D-E)** Axial CT image shows right subluxed facet (**D**, yellow ring) and naked facet sign (**E**, red arrow) and left laminar, and pedicle fracture (**E**, black arrow); **(F)** coronal reconstruction image shows the horizontal orientation of the laminar fracture (black arrow).

TLAOSIS based on injury morphology, neurological status, and PLC injury status (Table 1). TLAOSIS recommends conservative treatment if the score <4, surgical treatment if the score >5, and either surgical or conservative treatment ian the score of 4 or 5 (Gray zone).[17]

Statistical Analysis

We used SPSS version 20.0 software (IBM Corporation, Armonk, New York, USA) to conduct statistical analyses with a significance level of $P < .05$. The difference in the injury characteristic data between patients with and without change in AOSpine classification after MRI was examined using appropriate tests. The McNemar–bowker test was used to assess the difference in the proportions between the 3 main types of AO classification (A, B, C) and patients with TLAOSIS ≤3, 4-5 or ≥5 (indicating the conservative, gray zone, and surgical treatment, respectively).[3] We calculated confidence intervals (CIs) for proportions using Wilson CIs.[9] The default analysis determines the correct accuracy of CT in fracture classification against MRI as a gold standard using the proposed definitions for PLC injury this study: ≥ 2 CT findings and black stripe discontinuity. In comparison, the secondary analysis describes how accurate CT is in fracture classification if we used an alternative definition of PLC injury in CT as any positive CT findings or in MRI as high signal intensity. The results of sensitivity analysis were compared with default analysis to show the impact of CT or MRI criteria for PLC injury on reclassification rate by MRI[22] We applied Fleiss's kappa coefficient to assess the interreader agreement on the AO classification subtypes (A, B, C)—for CT and MR for 3 or 4 reviewers, respectively. We also used Cohen's kappa (k) statistic to assess the interareader agreement on the AO classification for CT and MRI. The kappa values can be interpreted as follows: ≤.2 = slight agreement, .21-.4 = fair agreement, .41-.6 = moderate agreement, .61-.8 = substantial agreement, .81-.99 = almost perfect agreement, and 1.0 = perfect agreement.[25]

Sample Size Calculation

We based the sample size calculation on the tables from Bujang et al.[26] With a PLC prevalence of 50% in our center, a null hypothesis (H0) of sensitivity for the current test of approximately 50% (the expected alternative hypothesis is to obtain a higher sensitivity of up to 70% with the combined test) and 80% power, the required sample is 98 per outcome group for a total of 196 patients.[26]

Results

Baseline Demographics and Injury Characteristics

The baseline characteristics of 244 patients enrolled in the study are presented in Table 2. There were 83.6% males and
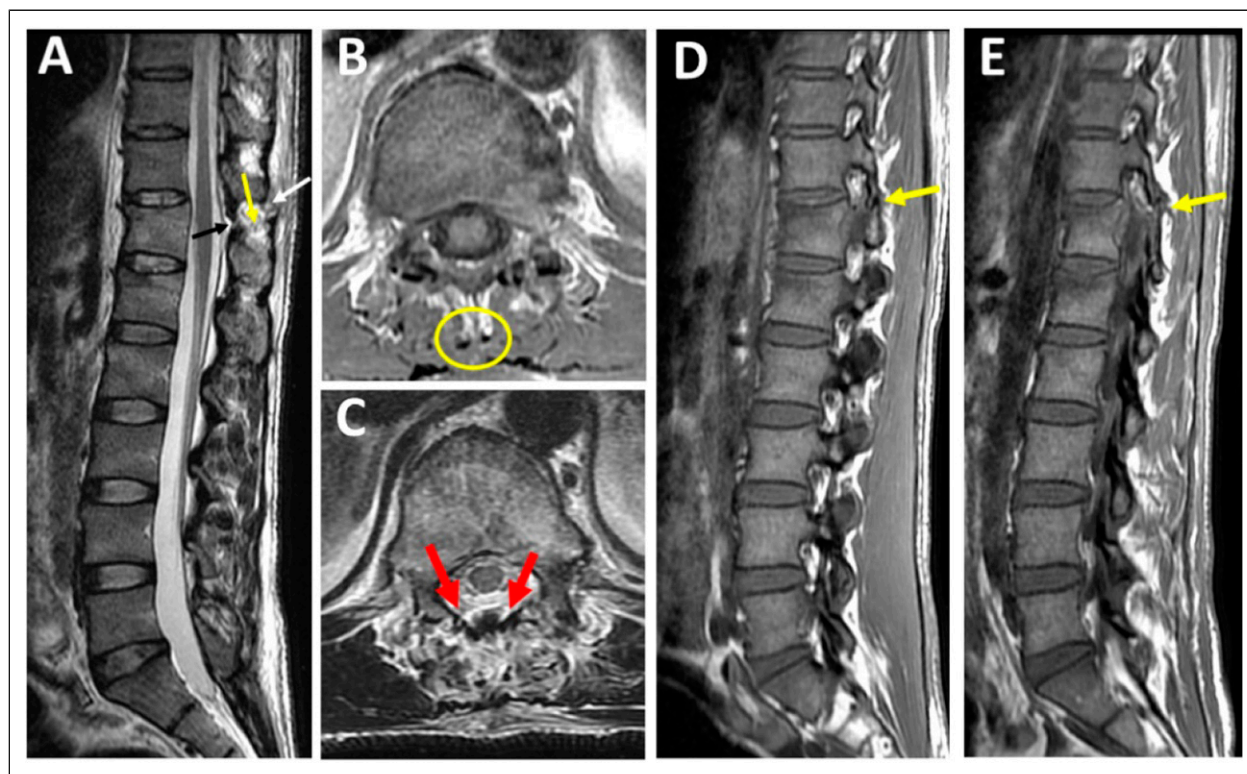
Figure 3. MRI Checklist for diagnosis of PLC injury. **(A)** Midsagittal T2-STIR shows black stripe discontinuity due to supraspinous ligament rupture (white arrow) or ligamentum flavum rupture (yellow arrow). High signal intensity due to interspinous ligament edema (black arrow), **(B-C)** Axial T2-STIR showing lack of signal intensity corresponding to a supraspinous ligament (B, yellow ring) or diffusely macerated ligamnetum flavum (C, 2 red arrows), **(D-E)** parasagittal T1-weighted images showing dislocated right and left facets (2 yellow arrows).

16.4% females, with a mean age of 31.5 ± 12 years. The most common mechanism of injury was a road traffic accident (73%) followed by a fall from a height of (20%). Fracture injury levels were thoracic, 18.5%; thoracolumbar, 68%; and low lumbar, 13.5%. Most MRIs (75%) were performed within 72 hours of injury. The main reason for the delayed MRI (3-10 days) was a referral from another hospital. Fifty-two percent of patients were treated surgically, while 48% were treated nonsurgically. Fifty percent of patients had multi-system injuries, while 32% had multi-level fractures.

Change in the Highest AO Classification After MRI

The highest AO classification by CT was type-A in 181 patients (74%), type-B in 54 patients (22%), and type-C in 9 patients (4%; Table 3). The distribution of A subtypes was A1 (25%), A2 (2.5%), A3 (19%), A4 (31%). The addition of MRI after CT changed the highest AO classification in 25/244 patients (10.2%, 95% CI [6.6%-15.3%], 1 proportion test $P < .0001$). The change in fracture classification by MRI was due to an upgrade from type A to type B in 20/244 patients (8.2%, 95% CI [5.0%-12.3%], $P < .0001$) and downgrade from type B to type A in 5/244 patients (2%, 95% CI [.64%-4.7%], $P < .11$). MRI did not change the AO classification in patients with Type-C fractures.

Change in the Thoracolumbar AOSpine Injury Score After MRI

The TL AOSIS by CT was ≤ 3 in 88 patients (36%), 4-5 in 74 patients (30%), and >5 in 82 patients (34%; Table 3). The overall McNemar-bowker test was significant indicating high agreement up to 86%. The addition of MRI changed TL AOSIS among the 3 treatment recommendation groups in 35/244(19.7%, 95% CI [14.9%-25.2%]. The addition of MRI changed TL AOSIS from ≤ 3 to >5 in 6/88 patients (6.8%, 95% CI [2.5%-14.2%], $P < .0001$), from >5 to ≤ 3 in 1/82 patients (1.2%, 95% CI [.04%-6.2%], $P = .85$) or >5 to 5 in 17/82 patients 20.7%, 95% CI [12.6%-31.0%, $P < .001$]. In patients with TL AOSIS 4 or 5, the addition of MRI changed TLAOSIS to >5 or <3 in 24/74 patients (32.4%, 95% CI [22%-44.3%], $P < .0001$).

Secondary (Sensitivity) Analysis

The default analysis (Figure 6A) shows that CT was accurate in fracture classification compared to MRI in 90% of cases, leading to a change in fracture classification by MRI in 10%. The results of the secondary analysis are described in Figure 6B and 6C. When using any positive CT findings as criterion for PLC Injury, CT's accuracy dropped to 68%, and
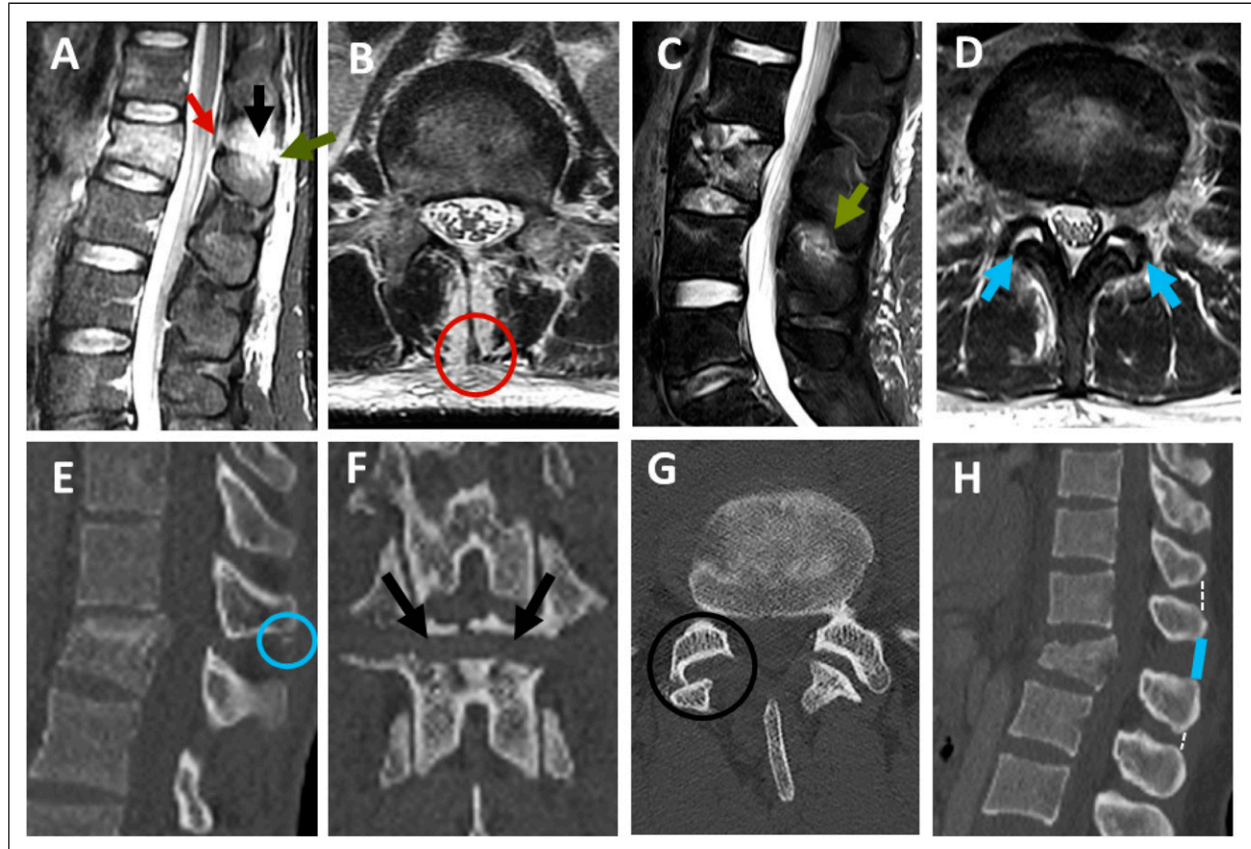
Figure 4. Grading of PLC injury in CT and MRI. **(A-D)** PLC integrity in MRI as graded as follows: **(A)** discontinuous black stripe due to disruption of the supraspinous ligament (red arrow) or igamentum flavum (green arrow); **(B)** disruption of the supraspinous ligament (red circle); **(C)** high-signal intensity due to interspinous ligament edema(green arrow); **(D)** high signal intensity due to facet joint effusion (2 green arrows). **(E-H)** CT findings suggestive of PLC injury; **(E)** avulsion or transverse fracture of the spinous process (blue circle); **(F)** horizontal fracture of the lamina or pedicle (2 black arrows); **(G)** facet joint malalignment (black circle); and **(H)** interspinous distance widening (ISW) >4 mm compared with the adjacent levels (solid line).

MRI changed classification by 32%, $P < .0001$). When using or MRI criterion as high signal intensity, the CT's accuracy dropped to 79%, and MRI changed classification by 21, % $P < .0001$).

Predictors of Change in TLFs Classification by MRI

As shown in Table 2, Thoracic fractures showed a significantly higher reclassification rate than thoracolumbar and low lumbar (20% vs 10% and 0%, respectively, $P = .07$). Fractures classified as type-B based on only 2 positive CT signs had a significantly higher rate of reclassification by MRI than those associated with 3 or 4 signs (17% vs 0%, $P = .03$, Table 2) As Table 4 shows, the rate of the upgrading to type-B for A subtypes A1-A4) was 8.6%, 16.6%, 9%, and 14%, respectively. The presence of M1 modifier increased the rate of upgrade to type B for all type A fractures (26% vs 4.6%, $P < .0001$), for A4 (32% vs 7%, $P < .006$), A3 (20% vs 0%, $P < .02$), but not for A2 (100% vs 0%, $P = .83$), and for A1(10% vs 8%, $P = .84$).

Long-Term Follow-up of Patients With Change AOSpone Classification After MRI

Of the 25 patients with fracture reclassification by MRI,19 (76%) were treated surgically, while the remaining 6 (24%, Figure 7) patients were treated non-operatively. Of the 6 patients treated conservatively, 1 required surgery for posttraumatic kyphosis, 3 had significant back pain, and two were symptom-free at the latest follow-up (20-45 months).

Inter- and Intrareader Agreement on CT and MRI Classification

The intrareader agreement on CT and MRI classification (AO type A vs B vs C) were almost perfect (Cohen k = .86, and .89). The inter-reader agreement on CT classification was almost perfect (Fleiss k = .84, whereas MRI classification was substantial (Fleiss k = .65).[25]
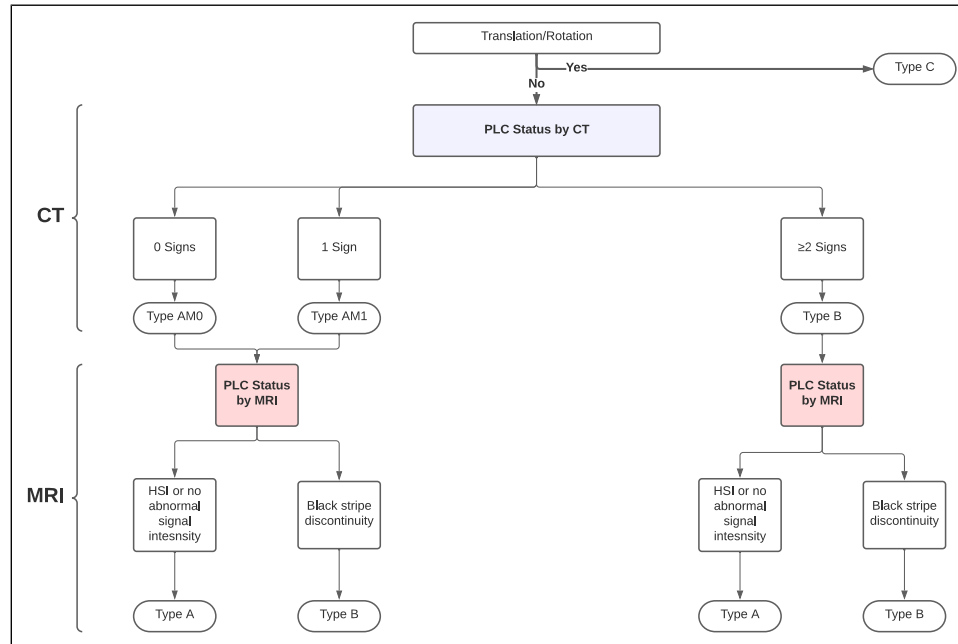
Figure 5. Algorithm for AOSpine classidcation in CT and MRI using our proposed criteria. Abbreviations: HSI, High signal intenisty

Discussion

To the best of our knowledge, this is the largest study (n = 244) to date that examines the impact of MRI on TLFs classification compared to CT alone.[7] This study concludes a relatively low rate of 10.2% for TLFs reclassification by MRI due to an upgrade from type A to B (8.2%) or downgrades from type B to A (2%).[8,9] MRI changed the TLOASIS among the 3 treatment's recommendation categories (≤3, 4-5, >5) in 48/244 (19.7%) patients. The presence of a single positive CT finding was the best predictor of an upgrade from type A to type B (26% vs 4.6%, $P < .0001$), whereas type-B fractures with only 2 positive CT findings, rather than 3 or more, was the best predictor for a downgrade from type B to type A (17% vs 0%, $P = .03$).

The significantly lower rate of TLFs' reclassification by MRI reported here (10.2% vs 30% previously) could be explained by the differences in sample size, patient population, and the CT/MRI criteria for PLC injury.[7,8] Our study is the first to analyze the impact of MRI on TLF's classification employing the new AO classification (rather than the Megrel classification), which is known for its simplicity and improved inter-reader reliability.[4,8,9] This study eliminated the previously observed MRI misclassification bias when PLC injury was identified on MRI using HSI, despite this criterion's low specificity for identifying PLC injuries during surgery.[12-21] Instead, we identified PLC injury in MRI as black stripe discontinuity caused by SSL or LF disruption, based on the improved specificity described by Pizones et al[8] for that criterion. The sensitivity analysis best highlights the magnitude of this bias, which shows a

significantly higher reclassification rate by MRI utilizing HSI vs black stripe discontinuity (33% vs11%).[10] This study's large sample size and consecutive recruitment (n = 244) resulted in a well-balanced cohort in terms of fracture morphology and conservative vs surgical patients.[18,19] We excluded patients with neurological deficits since they are universally treated surgically, as was the case in our study, regardless of their MRI fracture classification.[27] Additionally, patients with neurological deficits typically undergo MRI to document neural damage; these instances are overrepresented, resulting in a spectrum bias. The study's emphasis on neurologically intact patients eliminates such bias, which may be a crucial strength. We also excluded patients with osteoporotic fractures as defined by low-energy trauma and those who had an MRI more than 10 days after the injury since MRI sensitivity may decline after that point.[20,28]

Impact of MRI Reclassification on Decision-Making

MRI changed the TLOASIS among the 3 treatment's recommendation categories (≤3, 4-5, >5) in 48/244 (19.7%) patients. Patients with TLAOSIS ≤3 (conservative) had only a 7% chance of being recommended surgery after MR results. In patients with TLAOSIS >5 (Surgical), MRI altered the decision to conservative in 1.2% of patients and gray zone in 21% of patients. These findings demonstrate the reliability of TLAOSIS based on CT vs MRI in guiding decision-making for conservative and surgical groups. In the gray zone (TL AOSIS 4 or 5), MRI has the highest possibility of modifying the decision (33%) either to

Table 1. The Thoracolumbar AOSpine Injury Score (TL AOSIS) (Vaccaro et al., 2016).

Fracture pattern		Points
Type A-Compression fractures		
A0	Minimal injuries such as transverse process fractures	0
A1	Wedge compression	1
A2	Pincer compression injury	2
A3	Incomplete burst fracture: Fracture that only involves a single endplate	3
A4	Complete burst fracture: Fracture that involves both endplates	5
Type B-Tension band injuries		
B1	Osseous disruption of the tension band	5
B2	Posterior tension band injury including ligamentous injury	6
B3	Anterior tension band injury	7
Type C		
C	Translation/rotation injuries	8
Neurological status		
N0	Neurologically intact patient	0
N1	Resolved transient neurological symptoms	1
N2	Persistent radicular symptom	2
N3	Incomplete spinal cord or cauda equina injury	4
N4	Complete spinal cord injury	4
NX	Neurologic exam is unobtainable	3
Case specific modifiers		
M1	Fractures in which the status of the posterior ligamentous complex is unclear	1
M2	Patient-specific morbidities affect the treatment algorithm such as ankylosing spondylitis and polytrauma	0
Treatment recommendation	Conservative treatment	<4
	Conservative or surgical treatment	4,5
	Surgical treatment	>5

surgery (11%) or conservative treatment (22%). We believe that the reported 19.6% rate of change in treatment decisions based on TLAOSIS may overestimate the true impact of MRI in decision-making. For example, ruling out PLC injury by MRI in A4M1 or B (in 17 patients) reduces TLAOSIS from >5 to 5; however, we believe that most of these patients will require surgical treatment due to severe bone instability. Conversely, in case of mild instability and suspected PLC injury, CT (A3M1) ruling out or PLC injury by MRI may change decision-making. Severe bony instability can predict surgery independently and precludes the added value of change in TLAOSIS after MRI. TLAOSIS distinguishes the degree of bony instability between A4 and A3 fractures by assigning 5 points vs 3 points, respectively.[17] However, there is no consensus on clinically significant bony instability in burst fractures, and the A3/A4 criterion has been criticized for its inability to guide decision-making in all cases of burst fractures.[29,30] To assess the impact of Fracture reclassification after MRI on long-term outcomes, we looked at long-term follow-up from a patient with fracture reclassified by MRI. However, this was hampered because only 6/25 (24%) patients were treated non-operatively; 1 required surgery for post-traumatic kyphosis and 3 had significant back pain at the latest follow-up.

Predictors of Fracture Reclassification by MRI

We found that the number of positive CT findings and the fracture level were the best predictor of fracture reclassification by MRI. Type-B fractures with only 2 positive CT signs had a considerably higher rate of downgrading to type-A after MRI than those associated with 3 or 4 signs (17% vs 0%). Type-A3M1 and A4M1 had a considerably higher rate of upgrading to type-B after MRI than A3M0 and A4M0 (20% vs 0%, P < .02 and 32% vs 7%, P < .006, respectively). The association of the M1 modifier with the rate of an upgrade for type-B gives credibility to our definition based on a single CT finding. The findings that the M1 modifier can help guide decision-making in A3/A4 in the neurologically intact patient may be crucial considering the dispute on decision-making for this group of patients.[3] We also discovered that thoracic fractures had much higher reclassification rates than thoracolumbar or low lumbar fractures (20% vs 10% and 0%, respectively). A possible reason for the higher rate of thoracic fracture reclassification after MRI is the anatomical or biomechanical variations or the observed difficulties in recognizing PLC injury in MRI, possibly related to regional variations of SSL/ISL.[31-33] We think that osteoporosis may be associated with a higher rate of

Table 2. Baseline demographic and injury characteristics of the study patients.

	All patients	No change in classification after MRI	Change in classification after MRI	P-value
	(n = 244)	(n = 219)	(n = 25)	
Age, years (mean ± SD)	31.5 ±12	31 ± 11.8	36 ± 13	.03[a]
Male Gender	204 (83.6)	185 (84.5)	19 (76.0)	.26[c]
Mechanism of injury				.73[b]
Road traffic accident	177 (72.5)	161 (73.5)	16 (64.0)	
Fall from a height	49 (20.1)	43 (19.6)	6 (24)	
Suicidal jump	6 (2.5)	5 (2.3)	1 (4)	
Others	12 (4.9)	10 (4.6)	2 (8)	
*Level of injury**				.014[c]
Thoracic	45 (18.4)	36 (16.4)	9 (36)d	
Thoracolumbar	166 (68.1)	150 (68.5)	16 (64.0) d	
Lumbar	33 (13.5)	33 (15.1)	0 (0) d	
Number of fractures				.18[c]
Single level	166 (68)	150 (68.5)	16 (64.0)	
Multilevel contiguous	38 (15.6)	36 (16.4)	2 (8.0)	
Multilevel non-contiguous	40 (16.4)	33 (15.1)	7 (28.0)	
Polytrauma				.67[c]
No	121 (49.6)	110 (50.2)	11 (44.0)	
Yes	123 (50.4)	109 (49.8)	14 (56.0)	
*Treatment**				.03[c]
Surgical	127 (52.0)	108 (49.3)	19 (76)	
Conservative	117 (48.0)	111 (50.7)	6 (24)	
*AO classification by CT**				<.57[b]
Type-A (0 or 1 CT sign)	181 (74.2)	161 (73.5)	20 (80.0)	
0 Signs (M0)	131 (53.7)	123 (56.2)	8 (32)	
1 sign (M1)	51 (20.9)	39 (17.8)	12 (48)	
Type-B (2, 3, 4 CT signs)	55 (22.5)	50 (22.8)	5 (20.0)	<.001[c]
2 signs	30 (12.3)	25 (11.4)	5 (20)[e]	
≥ 3 signs	32 (13.1)	32 (14.6)	0 (0)[e]	
Type-C	8 (3.3)	8 (3.7)	0	
TLICS MRI score (mean ± SD)	3.7 ± 2.3	3.5 ± 2.3	5.4 ± 2.2	.001[a]
AOTLICS score (mean ± SD)	4.15 ± 2	4.15 ± 2	4.6 ± 1.9	.24[a]
Timing of MRI after injury				.691[c]
0-3 days	183 (75.0)	166 (75.8)	17 (68)	
3-10 days	61 (25.0)	53 (24.2)	8 (32)	
Strength of MRI				.66[c]
1.5 T	125 (51.2)	111 (50.7)	14 (56.0)	
3.0 T	119 (48.8)	108 (49.3)	11 (44.0)	

Abbreviations: ASIA, American Spinal Injury Association; CT, computed tomography; NS, not statistically significant (*P* > .05).
Data are number of patients, with the percentage in parentheses. *Significant difference in the injury characteristic data between patients with and without change in AOSpine classification after MRI.
[a]P-values were calculated with Student's t-test.
[b]P-values were calculated with Fisher's exact test.
[c]P-values were calculated with the chi-squared test.
[d]Pairwise comparison was done by partition Chi-squared tests for (thoracic 20% vs thoracolumbar 10%) and for (thoracolumbar 10% vs lumbar 0%) P-value was .07 for each.
[e]The proportions of 2 positive CT signs vs ≥3 were compared by chi-squared (17% vs 0%) P-value was .03.

fracture reclassification by MRI due to reduced accuracy of CT in diagnosing PLC injuries. However, this observation could not be confirmed because suspected osteoporosis could be confirmed with DEXA (dual-energy x-ray absorptiometry) in only a small proportion of patients.[13]

Table 3. Change in the AOSpine fracture type of TLAOSIS after addition of MRI compared to CT alone.

AO fracture type by CT	AO fracture type by MRI			Total	Agreement, %	P-Value*
	Type A (n = 166)	Type B (n = 69)	Type C (n = 9)			
Type A (n = 181)	161	20	0	181	90	<.001
Type B (n = 54)	5	49	0	54		
Type C (n = 9)	0	0	9	9		
Total	166	69	9	244		

TL AOSIS by CT	TL AOSIS by MRI			Total	Agreement, %	P-Value
	≤ 3 or more conservative (n = 99)	4-5 gray (n = 54)	> 5 surgery (n = 91)			
≥ 3 or more conservative (n = 88)	82	0	6	88	80	<.001
4-5 gray zone(n = 74)	16	50	8	74		
> 5 surgical (n = 82)	1	17	64	82		
Total	99	67	78	244		

Abbreviations: **TLAOSIS**, Thoracolumbar AOSpine injury score; **MRI**, Magnetic resonance imaging; **CT**, Computed Tomography.
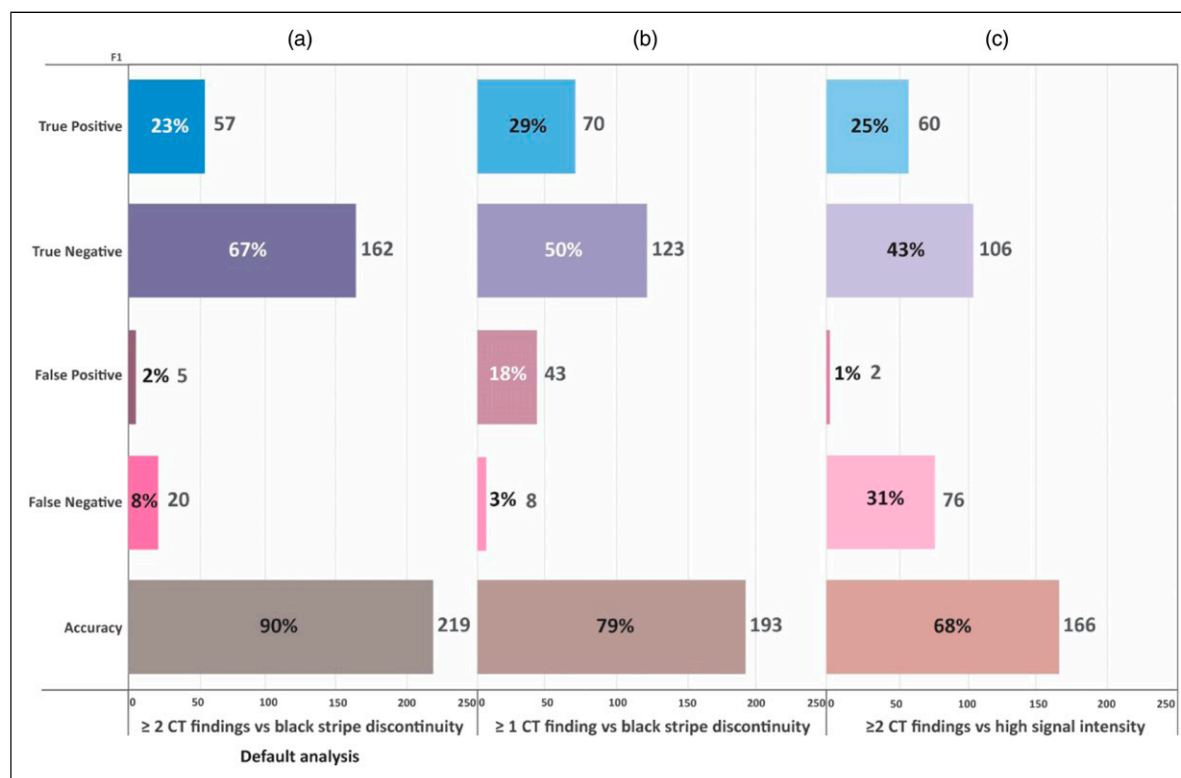*P-value was calculated using McNemar–bowker test.

Figure 6. Sensitivity analysis showing the impact of alternative CT/MRI criteria for PLC injury on the rate of fracture reclassification after MRI. **TP,** true positive; **TN,** True negative; **FP,** False positive; and **FN,** False negative. The default analysis **(A)** determines the correct accuracy of CT in fracture classification against MRI as a gold standard using the proposed criteria in this study: ≥ 2 CT findings and black stripe discontinuity. CT was accurate in fracture classification in 90% using these criteria, leading to a change in fracture classification by MRI in 10%. The secondary analysis **(B &C)** describes how accurate CT is in fracture classification if we used an alternative CT criterion as any positive CT findings instead of ≥2 CT findings (**B**, the CT's accuracy dropped to 68%, and MRI changed classification by 32%) or MRI criterion as high signal intensity (**C**, the CT's accuracy dropped to 79%, and MRI changed classification by 21%).

Table 4. Change the AO subtypes and Thoracolumbar AO injury severity Score after MRI compared to CT alone.

CT classification			MRI classification				
AOSpine	TLAOSIS	N	Change in AO classification, N (%)	Reason	Change in TL AOSIS, N (%)		P-value
TL AOSIS≤ 3 (non-operative treatment)		88	6/88 (6.8%)	Upgrade to type B	6/88 (6.8%)	≤3 to >5	
A1/M0	≤3	48	4/48 (8%)	Upgrade to type B	4/48 (8.3%)	≤3 to >5	0
A1/M1	≤3	10	1/10 (10%)	Upgrade to type B	1/10 (10%)	≤3 to >5	0
A2/M0	≤3	5	0/5 (0%)	Upgrade to type B	0/5 (0%)	≤3 to >5	0
A2/M1	≤3	1	1/1 (100%)	Upgrade to type B	1/1 (100%)	≤3 to >5	0
A3/M0	≤3	24	0/24 (0%)	Upgrade to type B	0/24 (0%)	≤3 to >5	0
TL AOSIS 4-5 (gray zone)	4- 5	74	8/74(10%)	Upgrade to type B	24/74(32%)	4- 5 > 5	
A4/M0	5	54	4/54 (7%)	Upgrade to type B	4/54 (7%)	5 to >5	0
A3/M1	4	20	4/20 (20%)	Upgrade to type B	20/20 (100%)	4 to >5 (20%) 4 to ≤3 (80%)	0
TL AOSIS >5 (operative treatment)	>5 to 5 or ≤3	82	12/82(14.6%)	Upgrade to type B or downgrade to A	5/82 (6%)	>5 to 5 or ≤3	
A4/M1	>5	19	6/19 (32%)	Upgrade to type B Downgrade to A4	6/19 (34%) 13/19(68%)	>5 to >5 >5 to 5	0
B	>5	54	4/54 (7%) 1/54(2%)	Downgrade to A4 Downgrade to A1	4/54 (7%) 1/54(2%)	>5 to 5 >5 to ≤3	0 0
C	>5	9	0/9(0%)	—	—	—	0
Total		244	25/244 (10.2%)	Upgrade to type B or down grade to A	48/244(19.7)	—	<.0001

Abbreviations: **TLAOSIS**, Thoracolumbar AOSpine Injury Severity Score; **CT**, Computed Tomography, **MRI**, Magnetic Resonance Imaging.

Implications

This study shows that TLFs without neurological deficit can be correctly classified with CT alone at a 10% misclassification rate.[7,8] The detailed fracture reclassification rate after MRI provided here for each fracture subtype or TLAOSIS category may help guide when to request MRI for neurologically intact patients with TLFs. Not having any patients with type C, A3M0, A2M0, and type B with ≥3 CT signs change classification after MRI assures us that these patients can be accurately classified with CT alone. In contrast, MRI might significantly change the fracture classification for A3 M1, A4 M0, or M1, B2 with 2 CT signs in 7-32% of patients, suggesting that MRI should be ordered for those patients. The above findings can substantially reduce the need for MRI in the emergency setting, reducing costs and efforts while also expediting treatment or in low-resource areas where MRI is unavailable.[16] The study emphasized the importance of using the appropriate CT criteria for PLC injury and a standardized approach of CT interpretation in achieving high CT accuracy for PLC injuries. Because there is no consensus on a legitimate definition of PLC injury in CT,[11,14] previous studies defined PLC injury in CT using any positive CT findings of PLC injury with no regard for the added value of combined CT findings.[9,8] Instead, we defined PLC injury on CT as ≥ 2 CT findings based on more recent studies.[16,15] When PLC injury on

CT was defined as ≥ 1 CT findings, our sensitivity analysis revealed a significantly higher reclassification rate by MRI than our default analysis (22% vs 11%). The standardized definitions for CT findings could explain the high inter-reader reliability on CT classification reported in that study (k = .84).[16,23,34] We previously proposed defining criteria for each CT finding based on examining the diagnostic value of various morphological features of CT findings for PLC injury.[16] For instance, we used a threshold of ISW> 4 mm, which was shown to predict PLC injury independently in MRI previously.[16] Furthermore, we developed a standardized imaging algorithm for PLC assessment, which includes the optimal imaging plan and strategies to address pitfalls in identifying each CT finding (Figure 2).[16,23]

Limitations

Our study has several limitations, including a retrospective, single-institution design requiring external validation through multicenter prospective studies.[18,19] Because MRI is an imperfect reference, intraoperative verification of PLC injuries may be warranted; however, that was not possible due to retrospective design.[35] Although intraoperative verification of PLC injuries is usually regarded as the gold standard, it is by definition restricted to surgical cases (in fact, open surgery), resulting in selection bias.[13] Notably, we excluded all patients with neurological deficits, including transient radiculopathy (N1) or persistent
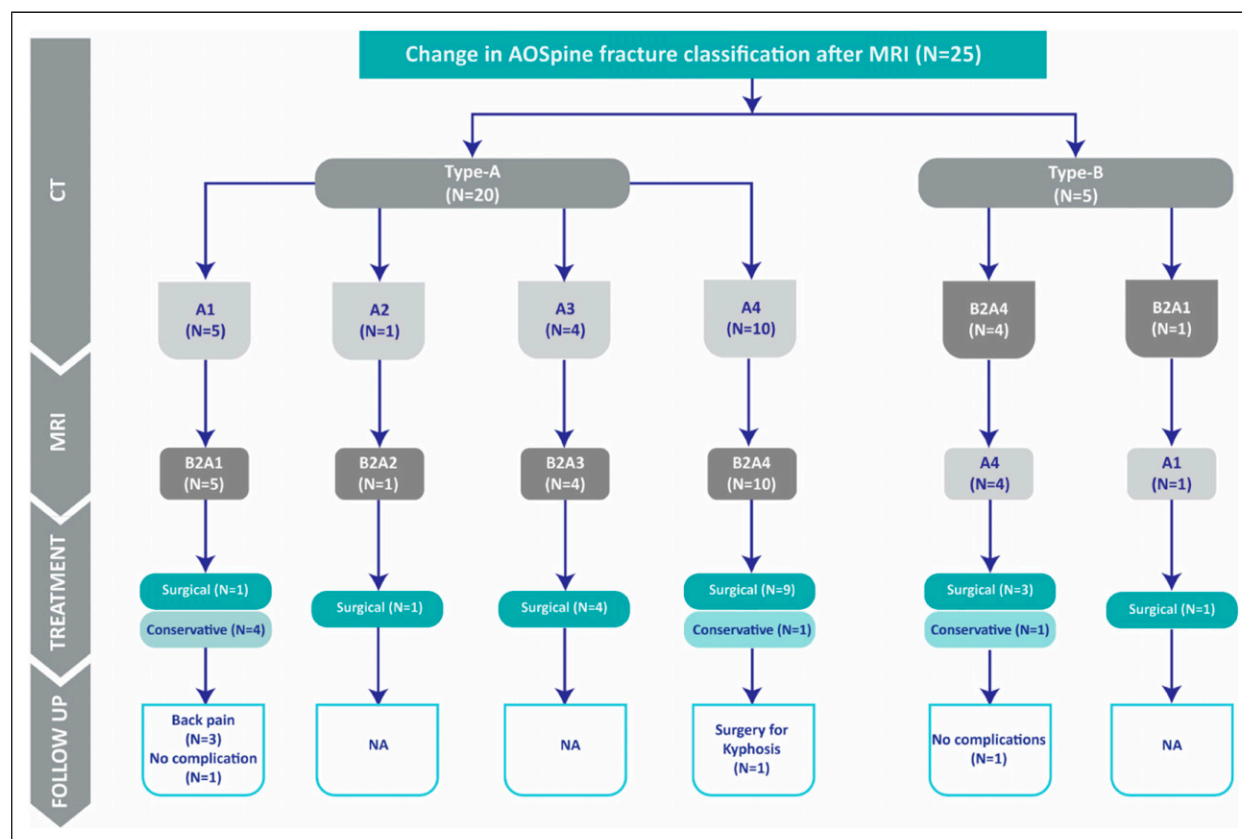
Figure 7. Follow-up data from 25 fractures with change in AOSpine fracture classification after MRI. **NA,** not applicable: follow-up data was limited to cases treated conservatively to observe sequela of missed PLC injuries.

radiculopathy (N2). Arguably, patients with N1 could be treated either surgically or conservatively. However, distinguishing between N1 and N2 was challenging because of the wide variability of presenting symptoms associated with N1 and retrospective design.[17] Finally, It should be noted that our findings were based on a consensus of 4 experienced reviewers, including a neuroradiologist, and that they may be less reproducible in less experienced hands or without prior training.[21,36]

Conclusions

This study concludes that using appropriate CT/MRI criteria of PLC injury, TLFs without neurological deficit can be correctly classified with CT alone at a 10% misclassification rate. When considering the algorithm suggested by TLAOSIS, MRI would have altered the decision-making in only 19.7% of patients compared to CT alone. The use of alternative CT/MRI criteria would significantly increase the rate of fracture reclassification up to 20-30%. The rate of change of fracture classification by MRI could be predicted by the number of positive CT findings on CT or fracture level (highest for thoracic and lowest for low lumbar).

ORCID iDs

Mohamed M. Aly https://orcid.org/0000-0001-5221-8288
Saleh Abduraba Ali https://orcid.org/0000-0002-9650-5796

Areej Al Fattani 🔘 https://orcid.org/0000-0002-9459-9936

References

1. Schroeder GD, Harrop JS, Vaccaro AR. Thoracolumbar trauma classification. *Neurosurg Clin N Am*. 2017;28(1):23-29. doi:10.1016/j.nec.2016.07.007.

2. Magerl F, Aebi M, Gertzbein SD, Harms J, Nazarian S. A comprehensive classification of thoracic and lumbar injuries. *Eur Spine J*. 1994;3(4):184-201. doi:10.1007/BF02221591.

3. Vaccaro AR, Lehman RA, 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(20):2325-2333. doi:10.1097/01.brs.0000182986.43345.cb.

4. Vaccaro AR, Oner C, Kepler CK, et al. AOSpine thoracolumbar spine injury classification system: fracture description, neurological status, and key modifiers. *Spine*. 2013;38(23): 2028-2037. doi:10.1097/BRS.0b013e3182a8a381.

5. Rajasekaran S, Vaccaro AR, Kanna RM, et al. The value of CT and MRI in the classification and surgical decision-making among spine surgeons in thoracolumbar spinal injuries. *Eur Spine J*. 2017;26(5):1463-1469. doi:10.1007/s00586-016-4623-0.

6. Leferink VJ, Veldhuis EF, Zimmerman KW, Ten Vergert EM, Ten Duis HJ. Classificational problems in ligamentary distraction type vertebral fractures: 30% of all B-type fractures are initially unrecognised. *Eur Spine J*. 2002;11(3):246-250. doi:10.1007/s00586-001-0366-6.

7. Qureshi S, Dhall SS, Anderson PA, et al. Congress of neurological surgeons systematic review and evidence-based guidelines on the evaluation and treatment of patients with thoracolumbar spine trauma: radiological evaluation. *Clin Neurosurg*. 2019;84(1):E28-E31. doi:10.1093/neuros/nyy373.

8. Pizones J, Izquierdo E, Alvarez 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. doi:10.1007/s00586-011-1913-4.

9. Winklhofer S, Thekkumthala-Sommer M, Schmidt D, et al. Magnetic resonance imaging frequently changes classification of acute traumatic thoracolumbar spine injuries. *Skeletal Radiol*. 2013;42(6):779-786. doi:10.1007/s00256-012-1551-x.

10. Schmidt RL, Factor RE. Understanding sources of bias in diagnostic accuracy studies. *Arch Pathol Lab Med*. 2013;137(4): 558-565. doi:10.5858/arpa.2012-0198-RA.

11. 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(3): 461-474. doi:10.1007/s00586-012-2602-7.

12. Vaccaro AR, Rihn JA, Saravanja D, et al. Injury of the posterior ligamentous complex of the thoracolumbar spine: a prospective evaluation of the diagnostic accuracy of magnetic resonance imaging. *Spine*. 2009;34(23):841-847. doi:10.1097/BRS.0b013e3181bd11be.

13. 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(9): 745-751.

14. Lee JY, Vaccaro AR, Schweitzer KM, et al. Assessment of injury to the thoracolumbar posterior ligamentous complex in the setting of normal-appearing plain radiography. *Spine J*. 2007; 7(4):422-427. doi:10.1016/j.spinee.2006.07.014.

15. Khurana B, Prevedello LM, Bono CM, et al. CT for thoracic and lumbar spine fractures: can CT findings accurately predict posterior ligament complex injury?. *Eur Spine J*. 2018;27(12): 3007-3015. doi:10.1007/s00586-018-5712-z.

16. Aly MM, Al-Shoaibi AM, Alzahrani AJ, Al Fattani A. Analysis of the combined computed tomography findings improves the accuracy of computed tomography for detecting posterior ligamentous complex injury of the thoracolumbar spine as defined by magnetic resonance imaging. *World Neurosurg*. 2021;151: e760. doi:10.1016/j.wneu.2021.04.106.

17. Vaccaro AR, Schroeder GD, Kepler CK, et al. The surgical algorithm for the AOSpine thoracolumbar spine injury classification system. *Eur Spine J*. 2016;25(4):1087-1094. doi:10.1007/s00586-015-3982-2.

18. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351(October):h5527. doi:10.1136/bmj.h5527.

19. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: arevised tool for quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(4):529-536.

20. 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. doi:10.3171/spi.2003.99.1.0020.

21. 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(8):747-753. doi:10.1016/j.spinee.2011.07.005.

22. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ*. 2013;346(May):f2778. doi:10.1136/bmj.f2778.

23. Aly MM, Al-Shoaibi AM, Al Fattani A, AlJuzair AH. Diagnostic value of various morphological features of horizontal and vertical laminar fractures for posterior ligamentous complex injury of the thoracolumbar spine as defined by MRI. *World Neurosurg*. 2021;153:e290-e299. Published online July 7. doi:10.1016/j.wneu.2021.06.109.

24. 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(11):2222-2231. doi:10.1007/s00586-012-2403-z.

25. Viera AJ, Garrett JM, Anthony J, Viera MD, JoanneGarrett M. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;3737(55):360360-363363. http://www1.cs.

columbia.edu/~julia/courses/CS6998/Interrater_agreement.
Kappa_statistic.pdf.

26. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. *J Clin Diagnostic Res*. 2016;10(10):YE01-YE06. doi:10.7860/JCDR/2016/18129.8744.

27. Vaccaro AR, Lim MR, Hurlbert RJ, et al. Surgical decision making for unstable thoracolumbar spine injuries. *J Spinal Disord Tech*. 2006;19(1):1-10. doi:10.1097/01. bsd.0000180080.59559.45.

28. Benedetti PF, Fahr LM, Kuhns LR, Hayman LA. MR imaging findings in spinal ligamentous injury. *Am J Roentgenol*. 2000; 175(3):661-665. doi:10.2214/ajr.175.3.1750661.

29. Lee JY, Vaccaro AR, Lim MR, 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. doi:10.1007/s00776-005-0956-y.

30. Aly MM, Al-Shoaibi AM, Al-Aithan A, et al. Can vertical laminar fracture further discriminate fracture severity between thoracolumbar AO type A3 and A4 fractures?. *World Neurosurg*. 2021;155:e177-e187. Published online August 14. doi:10.1016/J.WNEU.2021.08.035.

31. Moore TA, Bransford RJ, France JC, et al. Low lumbar fractures: does thoracolumbar injury classification and severity score work?. *Spine*. 2014;39(17):1021-1025. doi:10.1097/BRS. 0000000000000415.

32. Schroeder GD, Kepler CK, Koerner JD, et al. Can a thoracolumbar injury severity score be uniformly applied from T1 to L5 or are modifications necessary?. *Glob Spine J*. 2015;5(4): 339-345. doi:10.1055/s-0035-1549035.

33. Johnson GM, Zhang M. Regional differences within the human supraspinous and interspinous ligaments: a sheet plastination study. *Eur Spine J*. 2002;11(4):382-388. doi:10.1007/s00586-001-0378-2.

34. Maheswaran A, Aiyer SN, Farouk O, et al. Assessment of interobserver variability for CT scan–based evaluation of posterior ligament complex injury in thoracolumbar fractures: an international multicenter pilot study. *Glob Spine J*. 2020;10(2): 118-129. doi:10.1177/2192568219839414.

35. Vaccaro AR, Lee JY, Schweitzer KM, et al. Assessment of injury to the posterior ligamentous complex in thoracolumbar spine trauma. *Spine J*. 2006;6(5):524-528. doi:10.1016/j.spinee.2006. 01.017.

36. 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(4):391-396. doi:10. 3171/2009.10.SPINE08742.