The precision, accuracy and validity of detecting posterior ligamentous complex injuries of the thoracic and lumbar spine: a critical appraisal of the literature

Joost J. van Middendorp · Alpesh A. Patel ·
Michael Schuetz · Andrei F. Joaquim

Received: 8 June 2012 / Revised: 8 October 2012 / Accepted: 20 November 2012 / Published online: 1 December 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract

Purpose The diagnostic assessment and prognostic value of the posterior ligamentous complex (PLC) remains a controversial topic in the management of patients with thoracolumbar spinal injury. The purpose of this review was to critically appraise the literature and present an

overview of the: (1) precision, (2) accuracy, and (3) validity of detecting PLC injuries in patients with thoracic and lumbar spine trauma.

Methods Studies evaluating the precision, accuracy and/or validity of detecting and managing PLC injuries in patients with thoracic and/or lumbar spine injuries were searched through the Medline database (1966 to September 2011). References were retrieved and evaluated individually and independently by two authors.

Results Twenty-one eligible studies were identified. Few studies reported the use of countermeasures for sampling and measurement bias. In nine agreement studies, the PLC was assessed in various ways, ranging from use of booklets to a complete set of diagnostic imaging. Inter-rater and intra-rater kappa values ranged from 0.188 to 0.915 and 0.455 to 0.840, respectively. In nine accuracy studies, magnetic resonance (MR) imaging was most often ($n = 6$) compared with intra-operative findings. In general, MR imaging tended to demonstrate relatively high negative predictive values and relatively low positive predictive values for PLC injuries.

Conclusions A wide variety of methods have been applied in the evaluation of precision and accuracy of PLC injury detection, leaving spinal surgeons with a multitude of variable results. There is scant clinical evidence demonstrating the true prognostic value of detected PLC injuries in patients with thoracic and lumbar spine injuries. We recommend the conduct of longitudinal clinical follow-up studies on those cases assessed for precision and/or accuracy of PLC injuries.

Electronic supplementary material The online version of this article (doi:10.1007/s00586-012-2602-7) contains supplementary material, which is available to authorized users.

J. J. van Middendorp (✉)
Stoke Mandeville Spinal Foundation, National Spinal Injuries
Centre, Stoke Mandeville Hospital, Aylesbury, UK
e-mail: jvanmiddendorp@gmail.com

J. J. van Middendorp
Harris Manchester College, University of Oxford, Oxford, UK

J. J. van Middendorp · M. Schuetz
Institute of Health and Biomedical Innovation,
Queensland University of Technology,
Kelvin Grove, QLD, Australia
e-mail: m.schuetz@qut.edu.au

A. A. Patel · A. F. Joaquim
Department of Orthopaedic Surgery and Rehabilitation,
Loyola University, Chicago, IL, USA
e-mail: alpesh2@gmail.com

M. Schuetz
Trauma Service, Princess Alexandra Hospital,
Brisbane, QLD, Australia

A. F. Joaquim
Division of Neurosurgery, Department of Neurology,
State University of Campinas, Campinas,
SP 13083-970, Brazil
e-mail: andjoaquim@yahoo.com

Keywords Posterior ligamentous complex · Thoracic spine · Lumbar spine · Thoracolumbar spine · Spinal injuries · Systematic review

Introduction

Many classifications systems have been proposed to assist spinal surgeons in the treatment decision-making of thoracic and lumbar spine injuries [1]. In order to be useful, a spinal injury classification system should be reliable, accurate and clinically relevant [2]. As such, a classification should facilitate a surgeon in determining the ‘stability’ of the injured spine and its potential consequences.

In his classical paper on the ‘fractures, dislocations, and fracture-dislocations of the spine’, Sir Frank Holdsworth [3] was the first to note that posttraumatic spinal stability depends upon whether or not the posterior ligament complex (PLC) remained intact. Holdsworth considered the following four structures forming the PLC: the facet joint capsules (FJs), supraspinous ligament (SSL), interspinous ligament (ISL), and the ligamentum flavum (LF).

The pivotal role of the PLC in determining posttraumatic spinal instability has recently been reinforced by the Spinal Trauma Study Group (STSG) [4, 5]. In both the thoracolumbar injury severity scale (TLISS) and thoracolumbar injury classification and severity score (TLICS), the PLC forms one of the three key ‘injury characteristics’ [4, 5]. Despite the introduction of these novel classifications, controversy about the diagnostic assessment and prognostic value of the PLC in the management of spinal trauma patients remains.

The aim of this review was to critically appraise the literature and present an overview of: (1) the precision, (2) accuracy, and (3) validity of detecting PLC injuries in patients after thoracic and lumbar spine trauma. This enabled us, and would enable the reader, to judge the validity of previous reports and reconsider the role of PLC injuries in the management of spinal injuries. Furthermore, it identifies areas for improvement to be considered in future research initiatives.

Materials and methods

Search strategy and study eligibility criteria

In order to identify relevant articles on the precision and accuracy of detecting PLC injuries, we conducted a computerized search using the Medline (1966 to September 2011) database. The search terms used in Medline, PubMed interface, are presented in Fig. 1. The initial screening for eligibility included the following two criteria. The study had to consist of a clearly defined group of patients with thoracic and/or lumbar spine injuries. In addition, data on precision, accuracy and/or validity of detecting and managing PLC injuries had to be presented. Studies combining patients with cervical spine injuries with patients with

thoracic and/or lumbar spine injuries were excluded. Articles published in languages other than English and articles without an abstract were also excluded.

Data abstraction

Titles and abstracts were first screened; potentially relevant articles and reports were then retrieved and evaluated individually and independently by two authors (J.J.vM, A.F.J.). To ensure that no relevant studies were missed by the Medline search, we also performed a manual cross-reference search of the citations of each included article to obtain further relevant studies. In each phase, and in all cases, disagreement concerning inclusion of articles was resolved by discussion and consensus agreement.

Critical appraisal

Three different components for the evaluation of PLC injuries were addressed during critical appraisal of included articles: precision, accuracy and validity.

Precision, or reliability, is the extent to which repeated measurements under similar conditions of the same case agree with one other. Referring to the detection of PLC injuries, three main sources of measurement variation have been identified [6]. (1) *Observer variation* is the most frequently assessed type of variation and can be split into two components: inter-rater reliability and intra-rater reliability. Inter-rater reliability assesses the reliability, or agreement, of identifying a PLC injury when measured by different people under similar conditions. Intra-rater reliability assesses the reliability, or reproducibility, of identifying a PLC injury when measured more than once by the same rater. The other two sources of measurement variation; (2) *instrument variation* and (3) *subject variation*, are generally less commonly evaluated in spinal trauma imaging studies. However, when comparing different studies these aspects become more important.

The most straightforward method to assess inter- or intra-rater reliability is to evaluate percent agreement. A more popular method of assessing agreement is the use of the Cohen’s Kappa (κ) statistic. Unlike percent agreement, Cohen’s Kappa assesses the amount of agreement actually observed relative to the amount of agreement that would be expected by chance [7]. The Cohen’s Kappa statistic is traditionally applied for assessing the agreement on two nominal categories between two observers [8]. Kappa values can range from zero, representing no agreement beyond what would be expected by chance, to a value of 1.0, representing perfect agreement. Landis and Koch [9] have proposed the following categories for strength of agreement for the kappa coefficient: ≤ 0 = poor, 0.01–0.20 = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate,

Fig. 1 Syntax used to identify potentially relevant references in Medline using the PubMed interface (1966 to September 2011)

```
("ligaments"[MeSH Terms] OR "ligaments"[All Fields] OR "ligament"[All Fields] OR "ligamentum"[All Fields] OR "posterior tension band"[All Fields] OR "Posterior Ligamentous Complex"[All Fields] OR ("disco-ligamentous"[All Fields] AND "complex"[All Fields]) OR "discoligamentous complex"[All Fields] OR "zygapophyseal joint"[MeSH Terms] OR ("zygapophyseal"[All Fields] AND "joint"[All Fields]) OR "zygapophyseal joint"[All Fields] OR ("facet"[All Fields] AND "joint"[All Fields]) OR "facet joint"[All Fields])

AND

("spine"[MeSH Terms] OR "spine"[All Fields] OR "vertebrae"[All Fields] OR "spinal"[All Fields])

AND

("fractures, bone"[MeSH Terms] OR ("fractures"[All Fields] AND "bone"[All Fields]) OR "bone fractures"[All Fields] OR "fracture"[All Fields]) OR ("wounds and injuries"[MeSH Terms] OR ("wounds"[All Fields] AND "injuries"[All Fields]) OR "wounds and injuries"[All Fields] OR "injury"[All Fields])

AND

("thorax"[MeSH Terms] OR "thorax"[All Fields] OR "thoracic"[All Fields] OR "lumbosacral region"[MeSH Terms] OR ("lumbosacral"[All Fields] AND "region"[All Fields]) OR "lumbosacral region"[All Fields] OR "lumbar"[All Fields] OR "thoracolumbar"[All Fields])
```

0.61–0.80 = substantial, and 0.81–1 = almost perfect. Although this categorization is commonly applied, no consensus exists as to what constitutes an “acceptable” kappa coefficient. To analyse the reliability of data of an ordinal, interval, or ratio level of measurement, the use of the intraclass correlation coefficient (ICC) is preferred.

Accuracy is the degree to which the measurement actually represents what it is intended to represent. The accuracy of a test is best assessed by comparing it whenever possible to a ‘reference standard’ technique that is considered to accurately represent the truth. Insightful reviews on measures of diagnostic accuracy, including sensitivity, specificity and positive and negative predictive values (PPVs and NPVs) have been published previously [10, 11]. For the purpose of the current review accuracy values were extracted or, when necessary, re-calculated based on data reported in the original study.

The combination of accuracy and precision measures of detecting PLC injuries does not provide a clinician with enough information necessary for treatment decision-making. Identifying and measuring PLC injuries should also be clinically meaningful as can be expressed in terms of content, construct and prognostic validity.

Content validity examines how well a measurement represents all aspects of the phenomena under study. Content validity of a measurement is often established through subjective judgments, i.e. face validity, about whether the relevance and applicability of a diagnostic item seems reasonable [2, 12].

Construct validity refers to how well a measurement conforms to theoretical constructs. For instance, if a

specific PLC injury characteristic is theoretically believed to guide between two treatment options, a measure of this characteristic that has construct validity would show this guidance on the choice of treatment [2, 12].

Predictive validity, or prognostic validity, refers to the degree to which a variable can predict other outcomes of the same construct. For instance, if a specific PLC injury characteristic is theoretically believed to have an impact on the post-traumatic stability of the spine, a measure of this characteristic that has predictive validity would accurately predict the outcome of this injury when left untreated [12].

Apart from these general study features, included articles were also reviewed for the reporting of more specific methodological study characteristics. Based on items covered by both the STROBE statement [13] and the STARD initiative [14], we abstracted the reporting of following bias prone features: method of data retrieval, method of patient identification, method of agreement evaluation, method of accuracy evaluation, study setting, duration between assessments, eligibility criteria, descriptive details of assessed spinal injuries, observer details, definitions of PLC injuries and (if applicable) details of magnetic resonance (MR) imaging.

Results

Search and screening results

The computerized search strategy resulted in 801 citations. After screening of titles and abstracts, 45 remaining

potential eligible articles were obtained for full-text screening. Twenty-six articles were excluded during full-text screening. The majority of the irrelevant articles did not present agreement, accuracy or validity parameters or also evaluated cervical spine injury patients. Cross-referencing resulted in two additional citations. This resulted in a total number of 21 included articles of this review. Eight studies evaluated the agreement on PLC injury evaluation [15–22], another eight studies evaluated the accuracy of detecting PLC injuries [23–30] and one study evaluated both agreement and accuracy [31]. Four studies reported on different validity components of PLC injuries [32–35].

Precision of detecting PLC injuries

Study characteristics and methodology

Nine studies reported on the precision of PLC injury detection (Table 1). In seven studies, PLC injury detection was part of a spinal injury classification validation process. In these classification validation studies, diagnostic modalities provided to observers ranged from clinical history descriptions to MR imaging. The method of assessing the status of the PLC varied considerably; in two studies the use of a booklet was reported, one study reported the use of a ‘survey’, one study reported the use of a CD-ROM and in two classification validation studies the method of PLC evaluation was not reported. Koh et al. [15] reported their methods in detail: all observers assessed all included subjects’ medical records, X-rays, CT and MR images. Two studies focused on the agreement on PLC injury detection outside the scope of spinal injury classifications. In these two studies, MR was the single diagnostic modality provided to observers.

An overview of methodological properties of the agreement studies is presented in Table 2. Only one study reported the setting of its conduct (patients were recruited in a level 1 trauma centre). Two studies reported the recruitment of cases from a consecutive series and, in another study cases were selected based on indeterminate plain-film radiographs. The injury characteristics of included cases, including number, type and level of injury, were partially presented in four studies. The experience of observers was specified in two studies. Two out of nine studies reported details on MR tesla values, sequence types and timing of MR (see Web table 1). The utilization of a fat-suppressed T2 MR imaging sequence was reported in one study.

Definitions

Five studies reported and applied the same ‘four-structure’ definition of the PLC (see Table 1). The applied definition

of the PLC was not reported in four studies. All of the classification validation studies evaluated PLC injury as a broad entity only, except from Schweitzer et al. [16], who also evaluated diastasis of the facet joints on CT and ‘posterior’ oedema-like signal on MR imaging. In the two remaining studies [17, 31], ruptures of the SSL and ISL on MR imaging were assessed and considered representative of a PLC injury.

Agreement values

Table 3 presents the agreement values on the detection of PLC injuries as reported in and extracted from included studies. Eight studies evaluated the agreement on detecting an injury of the PLC as a broad entity. Unweighted inter-rater kappa values were presented in all studies and varied from 0.188 to 0.803. Intra-rater kappa values for PLC injury detection ranged from 0.455 to 0.810. Two studies evaluated the agreement on the detection of inter- and supraspinous ligament injuries separately. The inter-rater kappa values on detecting SSL and ISL injuries ranged from 0.630 to 0.690 and 0.631–0.915, respectively. For the same PLC structures, Dai et al. [17] reported intra-rater kappa values of 0.771–0.805 and 0.803–0.840, respectively. Schweitzer et al. [16] reported low inter-rater kappa values for the detection of diastasis of the facet joints on CT and ‘posterior’ oedema-like signal on MR imaging (see Table 3).

Accuracy of detecting PLC injuries

Study characteristics and methodology

Nine studies reported on the accuracy of detecting PLC injuries (Table 4). In six studies, the accuracy of PLC injury detection on MR imaging was assessed by applying intra-operative PLC evaluation as a reference standard. MR imaging was applied as a reference standard in three studies. Two studies compared ultrasound findings with MR imaging and/or operative findings.

Although no study reported the setting of the involved study centres, five studies documented a detailed overview of the numbers, types and levels of spinal injuries under study (see Table 5). Whereas eight studies reported the number of observers assessing the diagnostic test under study, three studies clearly stated whether the test was assessed by the same observer in all cases. The number of observers of the reference was reported in four studies. The blinding status of the test and reference observers was reported in eight and five studies, respectively. Two studies documented blinding of both test and reference observers. Except for the study by Vordemvenne et al. [30], none of the accuracy studies provided details about the experience of the observers.

Table 1 Study characteristics of included studies evaluating the agreement of PLC injury detection

Author	Data collection	Classification validation	Cases (n)	Observers (n)	Type of agreement evaluated		Details of reassessment			Diagnostic modalities ^a	Method agreement evaluation	PLC definition	PLC injury measurements
					Inter-rater agreement	Intra-rater agreement	Same cases	Reordered randomly	Time after first scoring				
Haba et al. [31]	Retro	–	35	3	Y	N	–	–	–	MR	All MR images	SSL, ISL, LF and FICs	PLC, SSL, ISL
Vaccaro et al. [21]	Pro	TLISS	71	5	Y	Y	Y	Y	>5 weeks	CH, PE, X-ray, CT, MR	CD-ROM	NR	PLC
Harrop et al. [18]	Retro	TLISS	56	30	Y	Y	Y	Y	3 months	CH, PE, X-ray, CT, MR	Booklet	NR	PLC
Ratliff et al. [20]	Retro	TLISS	56	28	Y	Y	Y	Y	3 months	CH, PE, X-ray, CT, MR	Booklet	SSL, ISL, LF and FICs	PLC
Patel et al. [19]	Pro	TLISS	71/25	21	Y	N	N ^b	–	7 months	CH, PE, X-ray, CT, MR	NR	NR	PLC
Whang et al. [22]	Pro	TLISS, TLICS	25	5	Y	N	Y ^b	Y	3 months	CH, PE, X-ray, CT, MR	NR	NR	PLC
Schweitzer et al. [16]	Retro	TLICS	13	34	Y	N	–	–	–	CT, MR	Survey (NOS)	SSL, ISL, LF and FICs	PLC, FJ diastasis, posterior oedema signal
Dai et al. [17]	Retro	–	61	3	Y	Y	Y	Y	6 months	MR	All MR images	SSL, ISL, LF and FICs	SSL, ISL
Koh et al. [15]	Retro	TLICS	114	3	Y	Y	Y	NR	4 weeks	CH, PE, X-ray, CT, MR	Medical records, all X-rays, CT and MR images	SSL, ISL, LF and FICs	PLC

Retro retrospective, *Pro* prospective, *TLISS* Thoracolumbar Injury Severity Score, *TLICS* Thoracolumbar Injury Classification System, *Y* yes, *N* no, *NR* not reported, *MR* magnetic resonance imaging, *CH* clinical history, *PE* physical examination, *CT* computed tomography, *NOS* not otherwise specified, *SSL* supraspinous ligament, *ISL* interspinous ligament, *LF* facet joint capsule, *PLC* posterior ligamentous complex, *FJ* facet joint

^a Provided to observers

^b Two inter-rater agreement evaluations

Table 2 Reporting of basic methodological aspects in studies evaluating the agreement on PLC injury detection

Author	Setting	Recruitment method	Eligibility criteria	Types of injury	Levels of injury	No. of spinal injuries	Experience observers
Haba et al. [31]	N	N	Y	Y	Y	N	N
Vaccaro et al. [21]	N	N	N	N	N	N	N
Harrop et al. [18]	N	N	N	N	N	N	N
Ratliff et al. [20]	N	N	N	N	N	N	N
Patel et al. [19]	N	Y	N	Y	N	N	Y
Whang et al. [22]	N	N	N	N	N	N	Y
Schweitzer et al. [16]	Y	Y	Y	Y	Y	N	N
Dai et al. [17]	N	Y	Y	Y	Y	Y	N
Koh et al. [15]	N	N	Y	N	N	N	N

Y yes, N no

Four out of nine studies reported details on MR imaging tesla values, sequence types and timing of MR imaging (see Web table 1). The utilization of a fat-suppressed T2 MR imaging sequence was reported in five studies.

Definitions

Seven studies reported and applied the same ‘four-structure’ definition of the PLC (see Table 4). Vaccaro et al. [24] reported an additional fifth structure of the PLC; the thoracolumbar fascia. In the study by Leferink et al. [26], only the inter- and supraspinous ligaments were considered to form the PLC. Despite comparable definition of the PLC among included studies, there was a variety of PLC components evaluated for accuracy. Six studies evaluated the accuracy of detecting PLC injury as a broad entity and six studies evaluated the accuracy of detecting injuries of specific PLC structures. Using ultrasound diagnostics, Vordemvenne et al. [30] also evaluated the presence of a hematoma below the SSL, which was considered as an indirect sign of PLC injury.

Accuracy values

Accuracy data on the detection of PLC injuries are presented in Table 6. Two studies evaluated the accuracy of detecting an injury of the PLC as a broad entity on MR imaging [30, 31]. Although detailed accuracy data were not presented, the PPVs and NPVs were high (>93 %). Except for ultrasound, other diagnostic tests showed much lower NPVs for PLC injury detection than MR imaging.

Using surgical procedures as a reference, five studies evaluated the accuracy of MR imaging for detecting SSL injuries. The accuracy ranged from 73.9 % in the study with the largest patient sample ($n = 42$) [24] to 100 % in a study with the smallest patient sample ($n = 5$) by Moon

et al. [27]. The latter study presented equally accurate values for ultrasound (100 %). Comparing CT and myelography findings with MR imaging reference, Petersilge et al. [28] found a low sensitivity (28.6 %) for the detection of SSL injuries. In general, less false positive SSL injuries were detected with ultrasound and CT when compared to MR imaging. For the detection of ISL injuries, comparable accuracy data were found as for SSL injuries.

Compared to inter- and supraspinous ligament injuries, MR imaging demonstrated lower sensitivity values for the detection of ligamentum flavum injuries [23–25, 27]. Moon et al. [27] demonstrated optimal specificity rates (100 %) for ligamentum flavum injury detection with both ultrasound and MR imaging using surgery as a reference. The accuracy of detecting facet joint capsule disruptions and thoracolumbar fascia injuries was evaluated by Vaccaro et al. [24] and found to be relatively low (69–79 %). Vordemvenne et al. [30] found low predictive values for detecting a hematoma below the supraspinous ligament (see Table 6).

Measuring and managing PLC injuries: face and predictive validity

Four studies evaluating aspects related to the validity of PLC injury measurement and management were retrieved. In a survey study among 28 members of the STSG, participants were asked to rank a list of various criteria potentially indicative of PLC injury [34]. Vertebral body translation as discerned on plain radiographs was ranked as the most important indicator for PLC disruption, followed by “interspinous spacing greater than that of level above or below” as discerned on anteroposterior plain radiograph and “diastasis of the facet joints” on CT. Applying a similar study design, the STSG subsequently asked participants to rank a list of various criteria potentially

Table 3 Previously published agreement data on detecting PLC injury in patients with thoracic and lumbar injuries

PLC component/ agreement type	Author	Subgroup	Cases (n)	Observers (n)	% Agreement	Kappa unweighted	Kappa weighted	ICC	Spearman ICC
<i>PLC injury</i>									
Inter	Haba et al. [31]		35	3	–	0.803	–	–	–
	Vaccaro et al. [21]		71	5	66.1	0.350	–	–	0.480
	Harrop et al. [18]		56	30	59.4	0.336	0.423		0.508
	Schweitzer et al. [16]		13	34	48.9	0.188	–	0.280	–
	Whang et al. [22]	(1) TLISS	25	5	–	0.520	–	–	0.737
		(2) TLICS	25	5	–	0.520	–	–	0.616
	Ratliff et al. [20]	(1) USA observers	56	15	62.1	0.373	0.464	–	0.551
		(2) Non-USA observers	56	13	60.6	0.336	0.425	–	0.503
	Patel et al. [19]	(1) Before TLISS implementation	71	21	–	0.202	–	–	0.327
		(2) After TLISS implementation	25	21	–	0.534	–	–	0.737
Intra	Koh et al. [15]		114	3	–	0.641	–	–	–
	Vaccaro et al. [21]		71	5	–	0.480	–	–	0.590
	Harrop et al. [18]		56	28	68.4	0.477	0.556	–	–
	Ratliff et al. [20]	(1) USA observers	56	15	67.4	0.455	0.519	–	0.578
		(2) Non-USA observers	56	13	69.7	0.507	0.610	–	0.685
Intra	Koh et al. [15]		114	3	–	0.810	–	–	–
<i>SSL injury</i>									
Inter	Haba et al. [31]		35	3	–	0.690	–	–	–
	Dai et al. [17]	Session 1	61	3	75.4	0.670	–	–	–
		Session 2	61	3	72.1	0.630	–	–	–
Intra	Dai et al. [17]		61	3	–	0.771–0.805 ^a	–	–	–
<i>ISL injury</i>									
Inter	Haba et al. [31]		35	3	–	0.915	–	–	–
	Dai et al. [17]	Session 1	61	3	75.4	0.672	–	–	–
		Session 2	61	3	72.1	0.631	–	–	–
Intra	Dai et al. [17]		61	3	–	0.803–0.840 ^a	–	–	–
<i>FJ diastasis</i>									
Inter	Schweitzer et al. [16]		13	34	70.5	0.395	–	0.430	–
<i>Posterior oedema signal</i>									
Inter	Schweitzer et al. [16]		13	34	70.5	0.280	–	0.350	–

PLC posterior ligamentous complex, SSL supraspinous ligament, ISL interspinous ligament, FJ facet joint, Inter inter-rater agreement, Intra intra-rater agreement

^a Range of values presented as average kappa value was not reported

indicative of PLC injury in the setting of normal plain radiographs [33]. This time “diastasis of the facet joints” on CT was ranked as most important indicator for PLC injury, followed by “posterior oedema (high signal intensity) in region of PLC elements” on T2 STIR or FAT SAT sagittal MR imaging and “Disrupted PLC components” on T1 sagittal MR imaging.

In another survey study conducted by the STSG, participants were asked to formulate their personal approaches to the management of different case scenarios with thoracolumbar injuries [35]. Independent of a patient’s neurological status, signs suggestive of PLC injury resulted in a shift towards posterior and circumferential surgical treatment strategies among the participants.

Table 4 Study characteristics of included studies evaluating the accuracy of PLC injury detection

Author	Data collection	Total cases/ LOIs (n)	Cases/ LOIs (n)	Test	Observers (n)	Blinding observers test	Reference	Observers (n)	Blinding observer reference	PLC: definition	PLC injury measurement
Petersilge et al. [28]	Retro	21/25	21/25	CT, or myelography, or both	2 ^b	Y	MR	2 ^b	N	SSL, ISL, LF and FJC _s	SSL
Terk et al. [23]	Pro	68/75	6/NR	MR	2 ^b	NR	OP	NR	NR	SSL, ISL, LF and FJC _s	SSL, ISL, LF
Lee et al. [25]	Pro	34/34	34/34	Palpation	NR	NR	OP	NR	NR	SSL, ISL, LF and FJC _s	PLC
Leferink et al. [26]	Retro	160/160	34/34	X-ray	NR	NR	OP	NR	NR		PLC
			34/34	MR	1 ^c	Y	OP	NR	NR		SSL, ISL, LF
Moon et al. [27]	Pro	12/12	160/160	X-ray and CT	NR	Y	OP	NR	NR	SSL and ISL	PLC
				US	1 ^c	N	MR	1 ^c	N	SSL, ISL, LF and FJC _s	SSL, ISL, LF
Haba et al. [31]	Retro	35/NR	5/5	US	1 ^c	Y	OP	NR	NR		SSL, ISL, LF
			5/5	MR	1 ^c	Y	OP	NR	NR		SSL, ISL, LF
Vaccaro et al. [24]	Pro	42/62	42/62	MR	1 ^c	Y	OP	NR	NR	SSL, ISL, LF and FJC _s	PLC, SSL, ISL
Vordemvenne et al. [30]	Pro	18/18	18/54 ^a	US	1	Y	OP	NR	Y	SSL, ISL, LF and FJC _s	SSL, ISL, LF, FJC, TLF
Pizones et al. [29]	Pro	30/59	12/12	MR	NR	Y	OP	NR	Y		PLC, SSL, ISL, HemSSL
			27/38	X-ray and CT	2 ^c	Y	MR	2 ^c	N	SSL, ISL, LF and FJC _s	PLC

Retro retrospective, *Pro* prospective, *LOI* level of injury, *NR* not reported, *MR* magnetic resonance imaging, *CT* computed tomography, *US* ultrasound, *Y* yes *N* no, *OP* operation/surgery, *SSL* supraspinous ligament, *ISL* interspinous ligament, *LF* ligamentum flavum, *FJC* facet joint capsule, *TLF* thoracolumbar fascia, *PLC* posterior ligamentous complex, *HemSSL* hematoma below the SSL

^a Uninjured adjacent levels were also evaluated

^b Not clear whether it was (were) the same observer(s) in all cases

^c Diagnosis reached by consensus

Table 5 Reporting of basic methodological aspects in studies evaluating the accuracy of PLC injury detection

Author	Setting	Recruitment method	Eligibility criteria	Types of injury	Levels of injury	No. of spinal injuries	Experience observers
Petersilge et al. [28]	N	N	Y	Y	Y	Y	N
Terk et al. [23]	N	N	N	Y	N	Y	N
Lee et al. [25]	N	N	N	Y	Y	Y	N
Leferink et al. [26]	N	N	Y	Y	N	Y	N
Moon et al. [27]	N	N	N	N	N	Y	N
Haba et al. [31]	N	N	Y	Y	Y	N	N
Vaccaro et al. [24]	N	Y	Y	Y	Y	Y	N
Vordemvenne et al. [30]	N	N	Y	Y	Y	Y	Y ^a
Pizones et al. [29]	N	Y	Y	Y	Y	Y	N

Y yes, N no

^a Only the experience of test observer was reported

One study evaluating the prognostic value and predictive validity of the PLC status was retrieved. Alanay et al. [32] performed a non-comparative, prospective, longitudinal analysis on 15 consecutive patients with a thoracolumbar burst fracture without neurological deficits. All patients underwent MR imaging indicating no injury of the PLC. After non-operative treatment and a minimum follow-up of 2 years, a loss of sagittal alignment was observed without having impact on patient-reported outcome measures. It was concluded that an intact PLC, as discerned on MR imaging, does not prevent loss of sagittal correction of the thoracolumbar spine. The loss of correction, however, was no shown to have an observable clinical impact on patients.

Discussion

This systematic review of the literature clearly exposes the controversy about the diagnostic assessment and prognostic value of the PLC in spinal trauma patients. A variety of methods have been applied to evaluate the precision and accuracy of detecting PLC injuries, leaving spinal surgeons with a multitude of variable reported results. This review included level 3b and 4 of evidence studies only and demonstrated the included studies' susceptibility to both sampling and measurement bias [36]. Moreover, most studies did not accurately describe which structures discerned on which diagnostic assessments were evaluated for precision or accuracy. Finally, and perhaps most importantly, there is scant clinical evidence demonstrating the prognostic value of detected PLC injuries in patients with thoracic and lumbar spine injuries.

The majority of included the studies did not report the method of patient recruitment and the use of eligibility

criteria in detail. Consistent recruitment of patients with use of justified, unambiguous eligibility criteria reduces the magnitude of sampling bias. Especially considering the setting of specialized tertiary spinal trauma centres, referral bias has a profound impact on a study's external validity. A study conducted in a general hospital most likely includes a relatively large proportion of patients with 'mild' spinal column injuries without PLC injury. Such a filtered, unbalanced patient sample complicates the interpretation and generalizability of presented agreement and accuracy values. To increase insight into the potential magnitude of sampling bias, we recommend authors to present relevant injury characteristics of all included patients and, ideally, also of those patients not meeting inclusion criteria. In addition, detailed reporting of which standardized diagnostic tests have been performed before and after the moment of patient recruitment will improve the reader's understanding on the magnitude of workup, or verification, bias; another important confounding element in studies assessing the accuracy of PLC injury detection [6].

Most of the studies evaluating the agreement on detecting PLC injuries were published on behalf of the STSG [16, 18–22]. In these classification validation studies, the assessment of the PLC status was part of the validation process of the TLISS [18–22] and TLICS [16, 22]. Taking a closer look at the original paper introducing the TLISS classification, it becomes clear that neither the PLC, nor an injury of this structure, has been defined in detail [5]. This was partly acknowledged with the introduction of its successor, the TLICS [21]. In their discussion, the authors state: "*Strict criteria regarding what constitutes a PLC injury radiographically... may improve accuracy when using the TLICS system.*" [21] Unfortunately, these 'strict criteria' were not provided for the TLICS either and have yet to be elucidated.

Table 6 Previously published accuracy data of detecting PLC injury in patients with thoracic and lumbar injuries

PLC component injury	Author	Reference	Test	Cases (n)	Observers (n; Ref/test)	Sensitivity	Specificity	PPV	NPV	Accuracy
PLC	Haba et al. [31]	OP	MR	35	NR/3	NR	NR	94.6	93.3	NR
	Vordemvenne et al. [30]	OP	MR	12	NR/NR	100	100	100	100	100
	Leferink et al. [26]	OP	X-ray and CT	160	NR/NR	53.6	60.4	80	30.6	55.3
	Lee et al. [25, 33]	OP	X-ray	34	NR/NR	66.7	66.7	95.2	16.7	66.7
	Vordemvenne et al. [30]	OP	US	54	NR/1	99	75	61.1	99.5	NR
	Lee et al. [25, 33]	OP	Palpation	34	NR/NR	52	66.7	92.9	14.3	53.6
	Pizones et al. [29]	MR	X-ray and CT	38	NR/NR	50	100	100	77.4	81.6
SSL	Haba et al. [31]	OP	MR	35	NR/3	89.4	92.3	NR	NR	90.5
	Terk et al. [23]	OP	MR	6	NR/2	100	NR	100	NR	NR
	Vaccaro et al. [24] ^a	OP	MR	42	1/1	85.7	56.6	75	71.4	73.9
	Lee et al. [25, 33]	OP	MR	34	NR/1	92.9	80	96.3	67.6	90.9
	Moon et al. [27]	OP	MR	5	NR/1	100	100	100	100	100
	Moon et al. [27]	OP	US	5	NR/1	100	100	100	100	100
	Vordemvenne et al. [30]	OP	US	54	NR/1	99	89.2	78.5	99.6	NR
	Moon et al. [27]	MR	US	12	1/1	100	100	100	100	100
	Petersilge et al. [28]	MR	CT, or myelography, or both	21	2/2	28.6	100	100	78.3	80
ISL	Haba et al. [31]	OP	MR	35	NR/3	98.5	87.2	NR	NR	94.3
	Terk et al. [23]	OP	MR	6	NR/2	100	NR	100	NR	NR
	Vaccaro et al. [24] ^a	OP	MR	42	1/1	90.6	46.7	78.4	70	76.6
	Lee et al. [25, 33]	OP	MR	34	NR/1	100	75	96.7	100	97
	Moon et al. [27]	OP	MR	5	NR/1	100	100	100	100	100
	Moon et al. [27]	OP	US	5	NR/1	100	100	100	100	100
	Vordemvenne et al. [30]	OP	US	54	NR/1	91.6	99.6	99	96.4	NR
	Moon et al. [30]	MR	US	12	1/1	83.3	100	100	85.7	91.7
LF	Terk et al. [23]	OP	MR	6	NR/2	100	NR	100	NR	NR

Table 6 continued

PLC component injury	Author	Reference	Test	Cases (n)	Observers (n; Ref/test)	Sensitivity	Specificity	PPV	NPV	Accuracy
	Vaccaro et al. [24] ^a	OP	MR	42	1/1	80.8	66.7	75	73.7	74.5
	Lee et al. [25, 33]	OP	MR	34	NR/1	85.7	88.5	66.7	95.8	87.9
	Moon et al. [27]	OP	MR	5	NR/1	100	100	100	100	100
	Moon et al. [27]	OP	US	5	NR/1	0	100	–	80	80
	Moon et al. [27]	MR	US	12	1/1	0	100	–	91.7	91.7
FJC	Vaccaro et al. [24] ^a	OP	MR	42	1/1	79.2–91.7	65.2	70.4–73.3	75.0–88.2	72.3–78.7
TLF	Vaccaro et al. [24] ^a	OP	MR	42	1/1	83.3	37.5	75	50	69.2
HemSSL	Vordemvenne et al. [30]	OP	US	54	NR/1	61.9	80	71.4	72.7	NR

Italicized accuracy values have been calculated based on data reported in the original manuscript. Range of values presented as average kappa value was not reported

PLC posterior ligamentous complex, SSL supraspinous ligament, ISL interspinous ligament, LF ligamentum flavum, FJC facet joint capsule, TLF thoracolumbar fascia, HemSSL hematoma below the SSL, OP operation/surgery, MR magnetic resonance imaging, CT computed tomography, US ultrasound, NR not reported, PPV positive predictive value, NPV negative predictive value

^a Disruption and incomplete disruption categories of PLC injury were collapsed into a single category to calculate accuracy measures

Given that no clear definitions of the PLC and criteria for PLC injury were provided, the next logical question follows: how was agreement on detecting a PLC injury then assessed? Table 1 provides the answers. There is a considerable methodological difference between detecting PLC injuries with use of ‘booklets’ [18, 20] and ‘all available X-rays, CT and MR images’ [15]. This, in combination with a low methodological quality of most of the TLISS validation studies (see Table 2) [18, 20–22] leaves the presented agreement values incomparable to results from other precision studies. Those studies that assessed the precision of detecting ISL and SSL injuries specifically demonstrated promising agreement values [17, 31]. However, as only three radiologists were involved in these studies [17, 31] the agreement among spine specialists still warrants further investigation. Additionally, the relative predictive value of the ISL and SSL as isolated structures may be less important compared to the broader definition of the PLC.

Most of the studies evaluating the accuracy of detecting PLC injuries compared MR imaging with intra-operative findings. Only two out of the seven studies that applied intra-operative findings as a reference standard reported blinding of reference observers (surgeons) with respect to test (imaging) findings. Moreover, for the detection of

specific ISL and SSL injuries, we found a tendency of studies with small sample sizes and low methodological quality demonstrating high accuracy values [23, 27] and studies with larger sample sizes and higher methodological quality demonstrating lower accuracy values [24, 25]. Vaccaro et al. [24] demonstrated low specificity values for detecting ISL and SSL injuries on MR imaging, indicating a large proportion of false positively identified injuries. Although ultrasound and other radiographic diagnostics demonstrated higher specificity values than MR imaging, the latter modality demonstrated less false-negative ISL and SSL injuries. These diagnostic accuracy data indicate that MR imaging results in an overestimation of the number of PLC injuries. Therefore, MR imaging may not yet have met the requirements of becoming a new reference standard for the detection of specific PLC injuries as has been hypothesized in previous literature [23, 37]. It should be noted, however, that as diagnostic imaging technology continuously evolves, previously reported reliability and accuracy findings for MR imaging may become outdated over time. New advances in imaging technologies may result in more accurate assessments of the PLC’s integrity and may also lead to improved precision and agreement among surgeons.

The statement on the limited role of current MR imaging techniques is supported by the survey study reported by Vaccaro et al. [34]. The top three ranked criteria potentially indicative of PLC injury did not comprise MR imaging criteria. However, in the light of the current review this is a remarkable finding as the majority of previously published accuracy data of detecting PLC injuries focuses on MR imaging with surgery being applied as a reference (see Table 6). The foremost explanation for the survey's finding can be sought in the contemporary diagnostic workup of spinal injuries. In contrast to plain radiographs and CT, MR imaging is not part of the current, routine diagnostic workup in the clearance of the thoracic and lumbar spine [38]. Surgeons naturally try to obtain as much as possible clinically relevant information from available diagnostic imaging (i.e. plain radiographs and CT) before requesting additional diagnostic assessments (i.e. MR imaging). In addition, the survey focuses more on the PPV of indicators for a PLC injury (e.g. presence of “diastasis of the facet joints”) rather than the NPV of indicators for such an injury (e.g. absence of “diastasis of the facet joints”). The authors of the survey study rightfully discussed that the diagnostic management becomes less straightforward when no signs of PLC injury can be detected on routine plain radiography or CT [34].

The succeeding survey study published by Lee et al. [33] indicates that surgeons consider “various characteristics” that can be assessed on MR imaging as more helpful in detecting a PLC injury when compared to plain radiographs and CT. The question which arises then is: do all patients with a thoracic and/or lumbar spine injury

without signs of PLC disruption on radiography or CT require MR imaging? From a pragmatic point of view the answer to this diagnostic workup question would be no. However, this review demonstrates that, from a scientific point of view, evidence to support this pragmatic approach is missing.

In fact, to our knowledge no study has been published presenting details on the prognostic value of PLC injuries related to radiographic, neurological and patient-reported outcomes. Alanay et al. [32] already demonstrated that the integrity of the anterior column has a more profound effect on radiological outcomes than what has previously been attributed to the PLC. Except from this reference, no other clinical study has analysed Holdsworth's hypothesis on the crucial role of the PLC for posttraumatic spinal stability. In contrast to the strong lack of prognostic data, the current review included various studies that evaluated the accuracy of PLC injury detection. However, as visualized in Fig. 2, determining the accuracy of PLC injury detection is less important than the assessment of its precision and prognostic value. While the PLC status, as discerned on CT, may not be as accurate or detailed as current MR imaging techniques, routine CT findings may demonstrate comparable or even stronger relations with treatment outcomes than MR imaging (i.e. better prognostic validity). Insight into the prognostic value of PLC injuries may provide much more clinically relevant information for a spinal surgeon than knowledge on whether current diagnostic techniques do reflect the truth accurately or not.

Ideally, future research should consider all three components of PLC injuries in thoracic and/or lumbar spine

Fig. 2 A schematic representation of assessing agreement, accuracy and clinical relevance of PLC injury detection in patients with thoracic and lumbar spine trauma, using two different diagnostic modalities (CT alone vs. CT and MRI)

injuries: precision, accuracy and clinical relevance. The inability to assess the integrity of the PLC status in non-operatively treated patients (see Fig. 2) can be countered by analysing the precision and prognostic validity of detecting PLC injuries. Longitudinal follow-up of those cases assessed for precision and/or accuracy of PLC injuries are therefore needed and would bring the debate on the management of PLC injuries to a more constructive level. Only then reliability and accuracy measures on the PLC status can be translated to valuable arguments in the prognostication and treatment decision-making of spinal injuries.

Acknowledgments We would like to thank Dr. Priyesh Dhoke and Dr. Hamish Deverall for their insightful discussions during the writing of this manuscript.

Conflict of interest No funds were received in support of this study. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript. The authors have no financial interest in the subject of this article. The manuscript submitted does not contain information about medical device(s).

References

- Mirza SK, Mirza AJ, Chapman JR, Anderson PA (2002) Classifications of thoracic and lumbar fractures: rationale and supporting data. *J Am Acad Orthop Surg*. 10(5):364–377
- van Middendorp JJ, Audige L, Hanson B, Chapman JR, Hosman AJ (2010) What should an ideal spinal injury classification system consist of? A methodological review and conceptual proposal for future classifications. *Eur Spine J* 19(8):1238–1249
- Holdsworth FW (1963) Fractures, dislocations, and fracture-dislocations of the spine. *J Bone Jt Surg Br* 45-B(1):6–20
- Vaccaro AR, Lehman RA Jr, Hurlbert RJ, Anderson PA, Harris M, Hedlund R et al (2005) 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)* 30(20):2325–2333
- Vaccaro AR, Zeiller SC, Hurlbert RJ, Anderson PA, Harris M, Hedlund R et al (2005) The Thoracolumbar Injury Severity Score: a proposed treatment algorithm. *J Spinal Disord Tech* 18(3):209–215
- Daly LE, Bourke GJ (2007) Interpretation and uses of medical statistics, 5th edn. Blackwell Science, Oxford, UK, pp 381–392
- Chapter 8: Using nonexperimental research. In: Bordens KS, Abbott BB (eds) (2011) Research design and methods: a process approach, 8th edn. McGraw-Hill, New York, NY
- MacLure M, Willett WC (1987) Misinterpretation and misuse of the kappa statistic. *A J Epidemiol* 126(2):161–169
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics*. 33(1):159–174
- Langlotz CP (2003) Fundamental measures of diagnostic examination performance: usefulness for clinical decision making and research. *Radiology* 228(1):3–9
- van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ (2009) Diagnostic methods I: sensitivity, specificity, and other measures of accuracy. *Kidney Int* 75(12):1257–1263
- Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB (2006) Designing clinical research, 3rd edn. Lippincott, Williams and Wilkins, Philadelphia, PA, pp 37–49
- von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 370(9596):1453–1457
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 326(7379):41–44
- Koh YD, Kim DJ, Koh YW (2010) Reliability and Validity of Thoracolumbar Injury Classification and Severity Score (TLICS). *Asian Spine J*. 4(2):109–117
- Schweitzer KM, Vaccaro AR, Harrop JS, Hurlbert J, Carrino JA, Rehtine GR et al (2007) Interrater reliability of identifying indicators of posterior ligamentous complex disruption when plain films are indeterminate in thoracolumbar injuries. *J Orthop Sci*. 12(5):437–442
- Dai LY, Ding WG, Wang XY, Jiang LS, Jiang SD, Xu HZ (2009) Assessment of ligamentous injury in patients with thoracolumbar burst fractures using MRI. *J Trauma* 66(6):1610–1615
- Harrop JS, Vaccaro AR, Hurlbert RJ, Wilsey JT, Baron EM, Shaffrey CI et al (2006) Intrarater and interrater reliability and validity in the assessment of the mechanism of injury and integrity of the posterior ligamentous complex: a novel injury severity scoring system for thoracolumbar injuries. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2005. *J Neurosurg Spine*. 4(2):118–122
- Patel AA, Vaccaro AR, Albert TJ, Hilibrand AS, Harrop JS, Anderson DG et al (2007) 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)* 32(3):E105–E110
- Ratliff J, Anand N, Vaccaro AR, Lim MR, Lee JY, Arnold P et al (2007) Regional variability in use of a novel assessment of thoracolumbar spine fractures: United States versus international surgeons. *World J Emerg Surg*. 2:24
- Vaccaro AR, Baron EM, Sanfilippo J, Jacoby S, Steuve J, Grossman E et al (2006) Reliability of a novel classification system for thoracolumbar injuries: the Thoracolumbar Injury Severity Score. *Spine (Phila Pa 1976)* 31(11 Suppl):S62–S69 discussion S104
- Wang PG, Vaccaro AR, Poelstra KA, Patel AA, Anderson DG, Albert TJ et al (2007) The influence of fracture mechanism and morphology on the reliability and validity of two novel Thoracolumbar Injury Classification Systems. *Spine (Phila Pa 1976)* 32(7):791–795
- Terk MR, Hume-Neal M, Fraipont M, Ahmadi J, Colletti PM (1997) Injury of the posterior ligament complex in patients with acute spinal trauma: evaluation by MR imaging. *AJR Am J Roentgenol* 168(6):1481–1486
- Vaccaro AR, Rihn JA, Saravanja D, Anderson DG, Hilibrand AS, Albert TJ et al (2009) 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)* 34(23):E841–E847
- Lee HM, Kim HS, Kim DJ, Suk KS, Park JO, Kim NH (2000) Reliability of magnetic resonance imaging in detecting posterior ligament complex injury in thoracolumbar spinal fractures. *Spine (Phila Pa 1976)* 25(16):2079–2084
- Leferink VJ, Veldhuis EF, Zimmerman KW, ten Vergert EM, ten Duis HJ (2002) Classificational problems in ligamentary distraction type vertebral fractures: 30 % of all B-type fractures are initially unrecognised. *Eur Spine J* 11(3):246–250

27. Moon SH, Park MS, Suk KS, Suh JS, Lee SH, Kim NH et al (2002) Feasibility of ultrasound examination in posterior ligament complex injury of thoracolumbar spine fracture. *Spine (Phila Pa 1976)* 27(19):2154–2158
28. Petersilge CA, Pathria MN, Emery SE, Masaryk TJ (1995) Thoracolumbar burst fractures: evaluation with MR imaging. *Radiology* 194(1):49–54
29. Pizones J, Izquierdo E, Álvarez P, Sánchez-Mariscal F, Zúñiga L, Chimenó P et al (2011) Impact of magnetic resonance imaging on decision making for thoracolumbar traumatic fracture diagnosis and treatment. *Eur Spine J* 20:390–396
30. Vordemvenne T, Hartensuer R, Lohrer L, Vieth V, Fuchs T, Raschke MJ (2009) Is there a way to diagnose spinal instability in acute burst fractures by performing ultrasound? *Eur Spine J* 18(7):964–971
31. Haba H, Taneichi H, Kotani Y, Terae S, Abe S, Yoshikawa H et al (2003) Diagnostic accuracy of magnetic resonance imaging for detecting posterior ligamentous complex injury associated with thoracic and lumbar fractures. *J Neurosurg* 99(1 Suppl): 20–26
32. Alanay A, Yazici M, Acaroglu E, Turhan E, Cila A, Surat A (2004) Course of nonsurgical management of burst fractures with intact posterior ligamentous complex: an MRI study. *Spine (Phila Pa 1976)* 29(21):2425–2431
33. Lee JY, Vaccaro AR, Schweitzer KM Jr, Lim MR, Baron EM, Rampersaud R et al (2007) Assessment of injury to the thoracolumbar posterior ligamentous complex in the setting of normal-appearing plain radiography. *Spine J* 7(4):422–427
34. Vaccaro AR, Lee JY, Schweitzer KM Jr, Lim MR, Baron EM, Oner FC et al (2006) Assessment of injury to the posterior ligamentous complex in thoracolumbar spine trauma. *Spine J* 6(5):524–528
35. Vaccaro AR, Lim MR, Hurlbert RJ, Lehman RA Jr, Harrop J, Fisher DC et al (2006) Surgical decision making for unstable thoracolumbar spine injuries: results of a consensus panel review by the Spine Trauma Study Group. *J Spinal Disord Tech* 19(1): 1–10
36. Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B et al (2009) Oxford Centre for Evidence-based Medicine—Levels of Evidence (March 2009). Available from <http://www.cebm.net/index.aspx?o=1025>. Cited July 2012
37. Oner FC, van Gils APG, Dhert WJA, Verbout AJ (1999) MRI findings of thoracolumbar spine fractures: a categorisation based on MRI examinations of 100 fractures. *Skeletal Radiol* 28(8): 433–443
38. Diaz JJ Jr, Cullinane DC, Altman DT, Bokhari F, Cheng JS, Como J et al (2007) Practice management guidelines for the screening of thoracolumbar spine fracture. *J Trauma* 63(3):709–718