

# INTEROBSERVER EVALUATION OF TLICS SYSTEM TO TREAT THORACOLUMBAR FRACTURES

AVALIAÇÃO INTEROBSERVADORES DO SISTEMA TLICS PARA O TRATAMENTO DAS FRATURAS TORACOLUMBARES

EVALUACIÓN ENTRE OBSERVADORES DEL SISTEMA TLICS PARA EL TRATAMIENTO DE FRATURAS TORACOLUMBARES

BERNARDO JOSÉ MOREIRA CHAVES<sup>1</sup>, LUIS EDUARDO CARELLI TEIXEIRA DA SILVA<sup>1</sup>, LUIS ANTONIO MEDEIROS MOLITERNO<sup>1</sup>, RENATO TAVARES<sup>1</sup>

## ABSTRACT

**Objective:** To evaluate the interobserver agreement regarding the TLICS Classification (Thoracolumbar Injury Classification and Severity Score). Furthermore, evaluate the reliability, analyzing the correlation between the treatment indicated by TLICS system (surgical or conservative) and the treatment indicated by each evaluator surgeon. **Methods:** Imaging tests and clinical data of 22 patients with thoracolumbar fractures were analyzed by eight spine surgeons, and two main analyzes were performed: the first compared the interobserver agreement related to TLICS and the second compared the agreement between the treatment indicated by TLICS classification (surgical or conservative) and treatment indicated by each surgeon - based on his personal experience and the preferred classification. **Results:** Using the parameters of Landis and Koch for interpretation of Kappa value, the interobserver agreement of TLICS classification was considered moderate in our study ( $K=0.6$ ). The agreement between the indications of treatment (surgical or conservative) dictated by the classification and the indication of each surgeon was considered excellent, with kappa value of 0.89. **Conclusion:** We believe that the classification is a good tool for the evaluation and the treatment indication in thoracolumbar fractures.

**Keywords:** Thoracic vertebrae; Lumbar vertebrae; Spinal fractures; Spinal cord injuries.

## RESUMO

**Objetivos:** Avaliar a concordância interobservadores com relação à Classificação TLICS (Thoracolumbar Injury Classification and Severity Score). Além disso, avaliar a confiabilidade, analisando a concordância entre o tratamento indicado pelo sistema TLICS (cirúrgico ou conservador) e o tratamento que cada cirurgião avaliador indicou. **Métodos:** Foram avaliados retrospectivamente exames de imagem e dados clínicos de 22 pacientes com fraturas toracolombares por oito cirurgiões de coluna vertebral. Duas análises principais foram realizadas: a primeira comparando a concordância do sistema TLICS interobservadores e a segunda comparando a concordância entre o tratamento indicado por essa classificação (cirúrgico ou conservador) e o tratamento que cada cirurgião indicou – com base em sua experiência pessoal e na classificação de escolha. **Resultados:** Utilizando os parâmetros de Landis e Koch para interpretação do valor de Kappa, a concordância interobservadores da classificação TLICS foi considerada moderada em nosso estudo, com valor de Kappa de 0,6. A concordância entre as indicações de tratamento (cirúrgico ou conservador) ditadas pela classificação e as indicações de cada cirurgião foi considerada excelente, com valor de Kappa de 0,89. **Conclusão:** Acreditamos ser a classificação analisada uma boa ferramenta para a avaliação e para a indicação do tratamento nas fraturas toracolombares.

**Descritores:** Vértebras torácicas; Vértebras lombares; Fraturas da coluna vertebral; Traumatismos da medula espinal.

## RESUMEN

**Objetivos:** Evaluar la concordancia entre observadores con respecto a clasificación TLICS (Thoracolumbar Injury Classification and Severity Score). Además, evaluar la confiabilidad, analizando la correlación entre el tratamiento indicado por el sistema de TLICS (quirúrgico o conservador) y el tratamiento que cada cirujano evaluador eligió. **Fueron evaluados de forma retrospectiva las imágenes y los datos clínicos de 22 pacientes con fracturas toracolombares por ocho cirujanos de columna vertebral. Se realizaron dos análisis principales: el primero comparó la concordancia del sistema TLICS y el segundo comparó la correlación entre el tratamiento indicado por esta clasificación (quirúrgico o conservador) y el tratamiento indicado por cada cirujano evaluador – en base a su experiencia personal y su clasificación de preferencia. Resultados:** Utilizando los parámetros de Landis y Koch para la interpretación del valor de Kappa, la concordancia entre observadores con respecto a la clasificación TLICS se consideró moderada en nuestro estudio, con un valor de kappa de 0,6. La concordancia entre las indicaciones de tratamiento (quirúrgico o conservador) dictadas por la clasificación y la indicación de cada cirujano se consideró excelente, con valor de Kappa de 0,89. **Conclusión:** Creemos que la clasificación se considera una buena herramienta para la evaluación y la indicación del tratamiento de las fracturas toracolombares.

**Descriptores:** Vértebras torácicas; Vértebras lumbares; Fracturas de la columna vertebral; Traumatismos de la médula espinal.

## INTRODUCTION

Thoracolumbar fractures can be serious and usually occur in young patients who are victims of high energy trauma.<sup>1</sup> Several classifications have been developed over the last 50 years to guide the treatment of these fractures, but few of them are easily reproducible, or validated prior to being promoted.<sup>2</sup> Most of

them take into account stability, the trauma mechanism, neurological and ligamentous lesions, the three-column concept, and the morphology of the fractures.<sup>3-4</sup> The most currently accepted and widely used classification is that of the AO group (*Arbeitsgemeinschaft für Osteosynthesefragen* [Association for the Study of Internal Fixation]).<sup>5</sup>

1. Instituto Nacional de Ortopedia e Traumatologia (INTO) [National Institute of Orthopedics and Traumatology], Rio de Janeiro, RJ, Brazil.

Study conducted in the Center for Diseases of the Spine (CDCOL), Instituto Nacional de Ortopedia e Traumatologia (INTO), Rio de Janeiro, RJ, Brazil.  
Correspondence: Rua: Senador Furtado, 61, apto. 1009, Rio de Janeiro, Brazil. 20270-021. bmoreirachaves@yahoo.com.br

The main objectives of a classification system are: to facilitate communication among professionals, to assist with documentation and medical research, and most importantly, to guide the treatment and prognosis. Therefore, it is useful when considering the severity of a lesion, and serves as a tool for the indication of treatment.<sup>2</sup>

In 2005, a new classification system, known as the Thoracolumbar Injury Classification and Severity Score (TLICS)<sup>6</sup>, was introduced by Alexander Vaccaro *et al* following a review of 127 cases of thoracolumbar spine fractures by 40 experienced surgeons from 15 different institutions.

This system is primarily based on three variables (the morphology of the fracture, the integrity of the posterior ligamentous complex, and the neurological profile), divided into subgroups that are assigned points depending on the level of involvement of each of these variables. The treatment proposal is influenced and guided by the sum of the points.

Just as in the AO group classification, the morphological analysis of the fracture is sub-classified into three different types, according to the principal mechanism of trauma and the radiological aspect of the lesion: compression, rotation-translation, and distraction fractures.

Compression fractures include: axial compression, compression-flexion, and lateral compression, and one point is assigned to this type of fracture morphology. If there is an associated burst component, an extra point is added. Thus, a burst fracture, characterized by an increase in the interpedicular distance is assigned a total of two points - one for the compression mechanism, and an additional point because it is a burst-type fracture.

Rotation-translation fractures should be assigned three points. In turn, distraction fractures, which include distraction-flexion and distraction-extension, should be assigned four points.

Unlike the AO classification, in which a rotation-translation injury is considered more serious (type C), the TLICS classification considers a distraction fracture (type B in the AO classification) to be more serious than a fracture of the rotation-translation mechanism, therefore receiving a higher score (four points). This characteristic is extremely important because when different components are combined in the same fracture, the most serious one should always be considered in the counting of points.

The posterior ligamentous complex (PLC), which includes the interspinous ligament, the supraspinous ligament, the ligamentum flavum, and the facet capsular ligaments, has the function of protecting the spinal cord from the forces of flexion, compression, distraction, and translation. Furthermore, it is known that once compromised, the PLC does not have great healing potential, therefore its integrity is one of the key points for defining the concept of spinal stability.

Once again, compared to the AO classification, in which the integrity of the PLC is inherent to the morphology of the fracture (i.e., type A fractures usually do not involve PLC lesions, while type B and type C fractures are considered to have impact to the PLC), the PLC is analyzed separately in the TLICS classification.

When the PLC is to be found to be intact, no point is assigned, and when it is affected, three points are assigned. A PLC lesion can also be classified as suspected or undetermined, in which case two points are assigned.

Magnetic resonance imaging (MRI) is a tool that can be used to determine the integrity of the PLC. In cases where it is not available, involvement of the PLC should be evaluated using indirect signs such as separation of the spinous processes, diastasis and subluxation of the facet joints, and vertebral translation or rotation. When the evidence is subtle, it should be classified as suspected or undetermined.<sup>6</sup>

When the patient has no neurological impairment (Frankel E), the neurological status is considered intact, and no point is assigned. When a neurological deficit is present, it should be sub-classified and assigned points according to the severity and urgency of the clinical profile presented, as follows: nerve root lesion (two points assigned), complete lesion of the spinal

cord or of the conus medullaris (also two points), incomplete lesion (three points), and lesion of the cauda equina (also three points).

An incomplete spinal cord lesion receives a higher score than a complete spinal cord lesion. Thus, according to the classification, incomplete lesions have higher chances of a surgical indication. This is probably due to the higher chance of neurological recovery in incomplete lesions with surgical treatment for decompression of the canal.

For patients who receive a total score of five points or more, surgical treatment is indicated. Conservative treatment is indicated for those who receive a total score of three or less. When the total is four points, other criteria must be analyzed, such as: collapse in kyphosis, lateral angulation, open fractures, multiple rib fractures, the inability to use orthoses, sternum fractures, CET, amputation, rheumatoid arthritis, ankylosing spondylitis, osteoporosis, age, obesity, and associated comorbidities.<sup>6</sup>

## METHODS

Following approval of the study by the Ethics Committee and one week prior to the evaluation, each of the eight surgeons participating in the study received a copy of the original article on the TLICS classification, and a summarized version in Portuguese.

Twenty-two cases of patients suffering from traumatic thoracolumbar fractures were randomly selected and projected via multimedia during a single session for all eight evaluators, all spine surgeons from our service. They were not allowed to communicate with each other during the entire presentation of the cases.

Each clinical case included the patient's age and sex, the type of accident, and the imaging exams, although not all of them had an MRI scan. The neurological profile from the preoperative evaluation of each patient, obtained from the medical records, was also projected for each clinical case, and was the only classification data provided. This effectively eliminated any bias of confusion, because if each evaluator were to examine each patient, we would also be analyzing, together with it, the interobserver agreement of the physical neurological exam among the evaluators.

On the day of the evaluation, after a brief presentation of the classification, each of the evaluators was provided with a new summary of the TLICS classification, and a simple questionnaire in which it was only necessary to check the sub-items and add up the points. This questionnaire (Attachment 1) also had spaces to mark the indication for treatment (whether conservative or surgical) based on the personal experience of each surgeon and on the classification (or classifications) they used in their clinical practice. Despite the fact that all the evaluators preferred to use the AO classification, we decided not to perform a direct comparative analysis (as has been done previously).<sup>5</sup> This meant the evaluators could also associate their own personal experience, as well as other theoretical concepts and other classifications.

Two principal analyses were conducted using the data obtained. The first was the evaluation of the TLICS system in terms of interobserver agreement of the classification itself. The second was the agreement between the treatment indicated by the TLICS classification (whether conservative treatment or surgery) and the treatment indicated by each surgeon (whether conservative treatment or surgery).

The data obtained was analyzed and the Kappa values were calculated, all of which were statistically significant, with a 95% confidence interval.

Using the Landis and Koch statistical parameters to interpret the Kappa values, the agreement is considered to be excellent when the Kappa value is greater than 0.75. When the value is between 0.4 and 0.75 the agreement is considered to be average to good. Values less than 0.4 are considered to represent poor agreement.

RESULTS

In the first analysis performed, regarding the interobserver agreement of the TLICS classification, the result obtained was moderate (Kappa = 0.60). (Chart 1)

In relation to the second analysis, in terms of the agreement between the treatment indicated by the classification and that indicated by each surgeon for each clinical case, the result was considered to be excellent (Kappa = 0.89). (Chart 2)

The interobserver Kappa value statistics for the TLICS classification were: average 0.60; minimum 0.37; maximum 0.84; mean 0.61; with a standard deviation of 0.12. The Kappa value statistics for the analysis of the agreement between the treatment indicated by the TLICS classification and that indicated by each surgeon, were: mean 0.89; minimum 0.61; maximum 1; median 1; with a standard deviation of 0.12.

In the second analysis, of the 176 evaluations performed (22 cases for each of the eight evaluators) in the study in question, eight (4.54%) had disagreement between the conduct suggested by the TLICS system and that based on the experience of each surgeon. The disagreements occurred in clinical case numbers 10, 12, and 14 (which did not include MRIs). Clinical case number 12, in which half of the evaluators opted for surgical treatment, stands out because the classification in question indicated conservative treatment.

Chart 1. Interobserver agreement of the TLICS classification.

Evaluator	A	B	C	D	E	F	G
A							
B	0.84						
C	0.47	0.57					
D	0.47	0.57	0.64				
E	0.50	0.47	0.52	0.62			
F	0.71	0.84	0.60	0.73	0.60		
G	0.67	0.81	0.71	0.70	0.37	0.67	
H	0.40	0.48	0.55	0.66	0.54	0.64	0.61

All the Kappa values referenced above were considered significant with a 95% confidence interval.

Chart 2. Interobserver agreement between the TLICS classification and the personal behavior of the evaluators.

Professional	Evaluator A	Evaluator B	Evaluator C	Evaluator D	Evaluator E	Evaluator F	Evaluator G
Evaluator A							
Evaluator B	0.78						
Evaluator C	0.78	1					
Evaluator D	0.78	1	1				
Evaluator E	0.61	0.78	0.78	0.78			
Evaluator F	0.78	1	1	1	0.78		
Evaluator G	0.78	1	1	1	0.78	1	
Evaluator H	0.78	1	1	1	0.78	1	1

All the Kappa values referenced above were considered significant with a 95% confidence interval.

DISCUSSION

The most widely accepted classification worldwide is the AO classification, and in spite of it also being the most used, it presents mild to moderate agreement results in most of the studies conducted.<sup>7,8</sup> Maçaneiro *et al*,<sup>7</sup> in a comparative study between the reproducibility of the AO and TLICS classifications, reported mild agreement for both classifications.

The TLISS system (Thoracolumbar Injury Severity Score)<sup>9</sup> is a slight variation of the TLICS system, published in the same year by the same author, in which he notes that the trauma mechanism is more important than the morphology of the fracture. And a small addition was included: in compression fractures, just as a burst fracture increases the score by a point, an angulation of more than 15 degrees in the coronal plane is assigned an additional point. Vaccaro *et al*,<sup>10</sup> in a study of seventy-one cases with five evaluators using the TLISS system, reported mild interobserver agreement for this classification.

The interobserver agreement for the TLICS/TLISS classification ranges from mild to moderate in the majority of studies,<sup>7,8,11,12</sup> but the interobserver agreement has presented good results.<sup>10,12</sup> Patel *et al*<sup>13</sup> demonstrated an improvement in the results of interobserver evaluations when reevaluated after seven months, suggesting that this is because the classification has achieved greater acceptance by the community of spine surgeons.

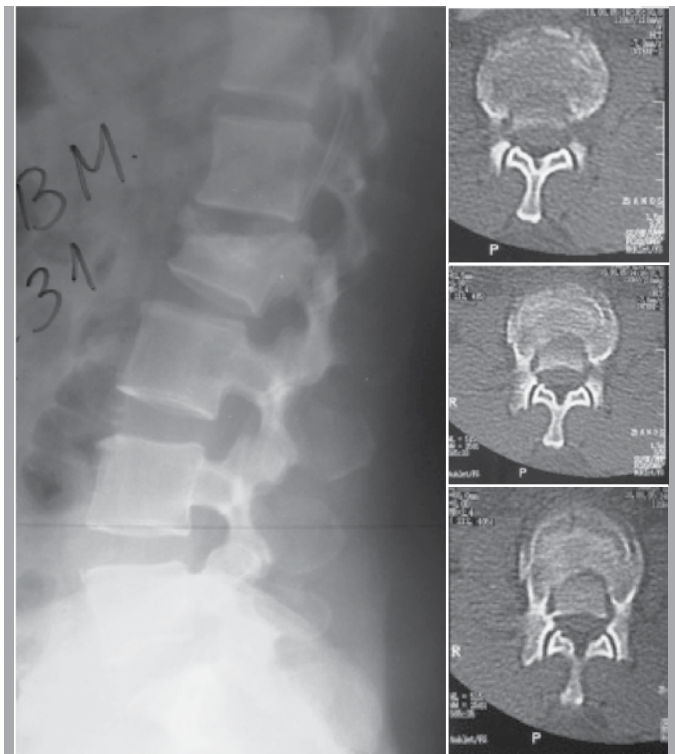
Our study corroborates the moderate results published in the literature for interobserver agreement of the TLICS/TLISS system. Our analysis also corroborates the results for the reliability of the TLICS/TLISS classification, as excellent results were obtained in regard to agreement between the treatments indicated by the classification and the treatment indicated by the evaluating surgeons, as previously demonstrated in various publications.<sup>10,7,8,12,14,15</sup>

We attribute the findings of our study to the fact that the new classification was well established before the case analysis. In addition to receiving the original article and a summary version in Portuguese a week earlier, on the day of the evaluation, the surgeons attended an oral presentation about the TLICS system and received an explanatory summary about it. We believe that the fact that it was only necessary to check off the sub-items on the questionnaire developed and distributed during the evaluation also contributed to the satisfactory results.

Several key points about the classification analyzed deserve comment. In the TLICS/TLISS classification, the integrity of the PLC is analyzed separately,<sup>16</sup> unlike the AO classification, in which the lesion is implicit in the morphology of the fracture, i.e., in type A fractures the integrity of the PLC is usually maintained and type B and C fractures usually present PLC injury due to the trauma mechanism. We believe that separate analysis of the PLC in the TLICS system offers a significant advantage because of its great importance to the concept of stability, and because it does not have the potential to heal in cases with conservative treatment. We should suspect PLC involvement whenever there is separation of the spinous processes, diastasis and subluxation of the facet joints, or vertebral translation or rotation. In these cases, magnetic resonance imaging (MRI) should be requested as a complementary exam.

Although the publication by Öner *et al*,<sup>11</sup> in their analysis of the AO classification, demonstrated no difference in interobserver agreement between a group of patients submitted to MRIs and another group without this exam, we believe that for the TLICS classification, the presence of the MRI can alter the results for intraobserver agreement. In our analysis, the three clinical cases without MRI were those that generated greater disagreement between the treatment indicated by the classification and the personal indication of each evaluator. Silva *et al*<sup>17</sup>, in their study of the TLICS system, also suggested that the presence of an MRI has a positive influence on the interobserver agreement of the classification.

The case that generated the greatest disagreement between the treatment indicated by the classification and that indicated by each surgeon was that of a 31-year-old patient without neurological deficit (Figure 1), for whom the TLICS system sugges-



**Figure 1.** Simple profile radiograph and CT axial cuts of a female patient, 31 years of age, victim of an automobile accident, without motor deficit in the physical exam, but with a large intracanal fragment seen in the imaging exams.

**REFERENCES**

- Hu R, Mustard CA, Burns C: Epidemiology of incident spinal fracture in a complete population. *Spine (Phila Pa 1976)*. 1996;21:492-499.
- Audigé L, Bhandari M, Hanson B, Kellam J. A concept for the validation of fracture classifications. *J Orthop Trauma*. 2000; 19(6):401-6.
- Denis F. The three column spine and its significance in the classification of acute thoracolumbar spinal injuries. *Spine (Phila Pa 1976)*. 1983; 8(8):817-31.
- Ferguson RL, Allen BL Jr. A mechanistic classification of thoracolumbar spine fractures. *Clin Orthop Relat Res*. 1984;(189):77-88.
- 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.
- Vaccaro AR, Lehman RA Jr, Hurlbert RJ, Anderson PA, Harris M, Hedlund R, et al. A new classification of thoracolumbar injuries: the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. *Spine (Phila Pa 1976)*. 2005;30(20):2325-33.
- Maçaneiro CA, Miyamoto RK, Lauffer RF, Larsen RV. Evaluation of the reproducibility between two classifications of fractures to the thoracolumbar spine and their correlations with treatment. *Coluna/Columna*. 2008;7(2):153-9.
- Wood KB, Khanna G, Vaccaro AR, Arnold PM, Harris MB, Mehbood AA. Assessment of two thoracolumbar fracture classification systems as used by multiple surgeons. *J Bone Joint Surg Am*. 2005; 87(7):1423-9.
- Vaccaro AR, Zeiller SC, Hulbert RJ, Anderson PA, Harris M, Hedlund R, Harrop J, Dvorak M, Wood K, Fehlings MG, Fisher C, Lehman RA Jr, Anderson DG, Bono CM, Kuklo T, Oner FC. The thoracolumbar injury severity score: a proposed treatment algorithm. *J Spinal Disord Tech*. 2005;18(3):209-15.

ted conservative treatment, while half of the evaluators indicated surgical treatment. This was due to an intercanal fragment larger than 50% of the canal, which was observed in the axial cuts in the Computed Tomography.

We take this opportunity to emphasize that other important points involved in the indication for surgical treatment must be taken into account, even though they are not directly present in the classification system. These include: the involvement and invasion of the spinal canal of more than fifty percent by the posterior wall, a decrease of more than fifty percent in vertebral height, progressive neurological deficit, and residual segmental kyphosis greater than 30 degrees<sup>18</sup>.

**CONCLUSION**

We believe that the TLICS system is reliable for the indication of treatment in thoracolumbar fractures given that our study contributed to the previously reported results about the excellent reliability of this classification. Regarding the validation of the classification, the interobserver agreement findings were moderate, which also is in line with earlier publications.

**ACKNOWLEDGEMENTS**

We thank the evaluators Antônio Eulálio, América Limoeiro, Luis Antônio Moliterno, Luis Eduardo Carelli Teixeira da Silva, Marcel Nascimento, Renato Tavares, Luiz Claudio Villela Schettino, and Ricardo Meirelles.

All authors declare no potential conflict of interest concerning this article.

**Attachment 1.** TLICS (Thoracolumbar Injury Classification and Severity Score) questionnaire prepared for the study, with the addition of a space for each surgeon's personal recommendation in their native language.

TLICS			
Neurologic Status		Posterior Ligament Complex (PLC)	
Intact	( ) 0 pnts	Intact	( ) 0 pnts
Nerve Root Injury	( ) 2 pnts	Suspected / Indetermined	( ) 2 pnts
Complete Spinal Cord (or Conus Medullaris) Injury	( ) 2 pnts	Disrupted	( ) 3 pnts
Incomplete Spinal Cord (or Conus Medullaris) Injury	( ) 3 pnts	<b>Based on TLICS</b>	
Cauda Equina	( ) 3 pnts	Score	
Morphology		Surgery ( ) Conservative ( ) Need to Evaluate Other Points ( )	
Compression	( ) 1pnt	<b>Based on Own Experience and Preferred Classification(s)</b>	
Associated Burst ?	( ) +1pnt	Surgery ( ) Conservative ( )	
Rotation / Translation	( ) 3 pnts		
Distraction	( ) 4 pnts		