RELIABILITY OF THE AO CLASSIFICATION OF THORACOLUMBAR FRACTURES COMPARED TO TLICS AND MAGERL

CONFIABILIDADE DA CLASSIFICAÇÃO AO DE FRATURAS TORACOLOMBARES EM COMPARAÇÃO COM TLICS E MAGERL

FIABILIDAD DE LA CLASIFICACIÓN AO DE FRACTURAS TORACOLUMBARES EN COMPARACIÓN CON TLICS Y MAGERL

Carlos Augusto Costa Marques[1,2], Xavier Soler Graells[1,2], Alynson Larocca Kulcheski[1,2], Gustavo Meurer[1,2], Marcel Benato[1,2], Pedro Grein Santoro[1,2]

1. Hospital do Trabalhador, Spine Surgery, Curitiba, PR, Brazil.
2. Universidade Federal do Paraná (UFPR), Hospital das Clínicas, Department of Orthopedics and Traumatology, Curitiba, PR, Brazil.

ABSTRACT

Objective: To test the reliability of the new AO/2013 classification compared with AO/Magerl and TLICS. Methods: Four spine surgeons retrospectively and blindly evaluated imaging and clinical data from 98 patients with thoracolumbar fractures. Results: Using the Kappa coefficient, we obtained the best reproducibility for the AO/2013 classification compared to the other two, represented by Kappa coefficient of 0.690. We could also obtain, with good reproducibility among the evaluators (Kappa 0.690), the most common subtypes of AO/2013 classification with indication for surgery. Conclusion: We believe that the new AO/2013 classification has proven to be a good communication tool among spine surgeons with good reproducibility, but more studies should be conducted in several centers in order to be consolidated and so that the prognosis between the types of injury is better understood.

Keywords: Spinal fractures; Spine; Lumbar vertebrae; Thoracic vertebrae; Classification.

RESUMO

Objetivos: Testar a confiabilidade da nova classificação AO/2013 com relação às classificações AO/Magerl e TLICS. Métodos: Foram avaliados retrospectivamente exames de imagem e dados clínicos de 98 pacientes com fraturas toracolombares por quatro cirurgiões de coluna vertebral, de modo cego. Resultados: Utilizando o coeficiente Kappa, obtivemos a melhor reprodutibilidade para a classificação AO/2013 quando comparada com as outras duas, representada por um índice Kappa de 0,690. Pudemos obter também, com boa reprodutibilidade entre os avaliadores (Kappa 0,690), os subtipos mais comuns da classificação AO/2013 com indicação de cirurgia. Conclusão: Acreditamos que a nova classificação AO/2013 demonstrou ser uma ótima ferramenta de comunicação entre os cirurgiões de coluna, com boa reprodutibilidade, porém mais estudos devem ser realizados em diversos centros para que seja consolidada e que o prognóstico entre os tipos de lesão seja mais bem compreendido.

Descritores: Fraturas da coluna vertebral; Coluna vertebral; Vértebras lombares; Vértebras torácicas; Classificação.

RESUMEN

Objetivos: Comprobar la fiabilidad de la nueva clasificación AO/2013 en relación con las clasificaciones AO/Magerl y TLICS. Métodos: Se evaluaron de forma retrospectiva imágenes y datos clínicos de 98 pacientes con fracturas toracolumbares por cuatro cirujanos de columna, a ciegas. Resultados: Utilizando el coeficiente Kappa, se obtuvo la mejor reproducibilidad para la clasificación AO/2013 en comparación con las otras dos, representada por un índice Kappa de 0,690. Hemos sido capaces de obtener también, con buena reproducibilidad entre evaluadores (Kappa 0,690), los subtipos más comunes de la clasificación AO/2013 con indicación para cirugía. Conclusión: Creemos que la nueva clasificación AO/2013 resultó ser una gran herramienta de comunicación entre los cirujanos de columna, con buena reproducibilidad, pero más estudios deben llevarse a cabo en varios centros para que se consolide y que el pronóstico entre los tipos de lesiones sea más bien comprendido.

Descriptores: Fracturas de la columna vertebral; Columna vertebral; Vértebras lumbares; Vértebras torácicas; Clasificación.

INTRODUCTION

As a consequence of the advances in the initial treatment of polytraumatized patients, more and more victims of serious spine injuries are showing up at emergency care units alive and needing immediate treatment.[1]

The age group most frequently affected by spine injuries is the one with the highest productivity for society, i.e., from 20 to 59 years of age.[1] The region between the T12 and L2 vertebrae is the site of more than 50% of all spine fractures. (Figures 1 and 2)

The attention given to spine fractures in Brazil has increased in the last few years, with the rise in urban violence and high-energy traumas, traffic accidents, and falls from heights, in addition to

Study conducted at the Hospital Cristo Redentor, Neurosurgery Service, Porto Alegre, RS, Brazil.
Correspondence: Rua Buenos Aires 100, Bairro Jardim das Américas, Cuiabá, MT, Brasil. 78060-634. ctauto@gmail.com

Received on 04/19/2016, accepted on 08/26/2016. **http://dx.doi.org/10.1590/S1808-185120171601162779**

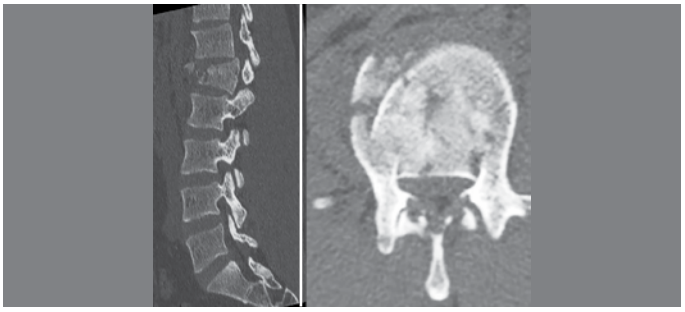

Figure 1. Axial computed tomography in sagittal and axial slices showing a type A4 fracture of vertebra L1.

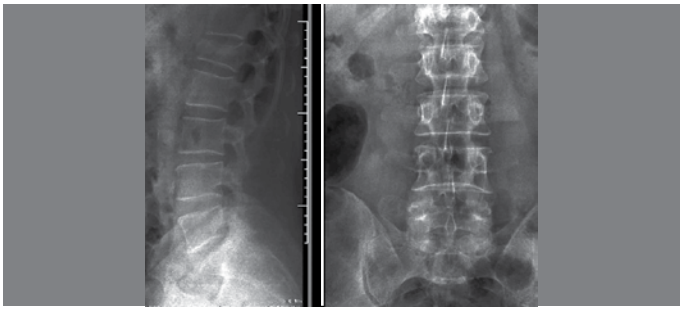

Figure 2. Anteroposterior and lateral radiography showing a type A4 fracture of vertebra L1.

the potential for neurological damage, which approaches 40% in cervical fractures and 15-20% in thoracic fractures.[2]

With the goal of unifying communication for better treatment of spine fractures, and in particular those of the thoracolumbar segment, several classification systems have been developed.[3]

In terms of the classifications based on pathomorphology, that of Magerl et al.[4] is the most detailed, with a total of 53 types of fractures. Up until now, this classification system has been used as an international reference. However, clinical application of this classification has been neither validated nor revised.[5,6] Watson-Jones[7] has already stated that the concept of instability would be critical to any algorithm regarding thoracolumbar fractures.

The classification known as TLICS (Thoracolumbar Injury Classification System)[8] assesses neurological status, the integrity of the posterior ligamentous complex (PLC), and the morphology of the lesion using descriptive categories.[8] It uses Magnetic Resonance Imaging (MRI) to evaluate the integrity of the PLC, a fact that has been evaluated in other studies.[9,10] We should note the importance of this classification, as it emerged in order to overcome several deficiencies of the AO/Magerl 1994 classification system.[4] With the recent publication of the new AO classification in 2013,[3] it needed to be validated in the medical community, and especially among the professionals who deal with this type of injury. The new AO/2013 classification maintains the group format (A, B, C) and the subtypes (A0, A1, A2, A3, A4, B1, B2, B3, and C).[3]

The objective of this study was to test the reliability of the new AO/2013 classification in our service, comparing it to the two previously used classification systems, AO/Magerl 1994[4] and TLICS.[8]

The research was developed at the Hospital do Trabalhador-UFPR, located in Curitiba-PR, which treats around 60% of the traumas of the capital and the metropolitan region.

METHODOLOGY

The study was approved by the Institutional Review Board as number CAAE: 42605915.5.0000.5225.

This was an observational, longitudinal, retrospective, and descriptive study in which we reviewed the medical reports of 100 cases of thoracolumbar spine fractures treated at the Hospital do Trabalhador-UFPR in Curitiba-PR, Brazil, during the period from January, 2013, to December, 2014. The following parameters were used for case selection:

Inclusion criteria: Patients with vertebral fractures at the T1 to L5 level, radiographs and computed axial tomography (CAT) taken at hospital admission.

Exclusion criteria: Patients with pathological fractures, incomplete medical records, fractures not at the proposed levels, inadequate imaging exams, firearm projectile fractures (FAP).

For each case we evaluated the radiographic images in anteroposterior (AP) and lateral orthogonal views and CAT images in coronal, sagittal, reconstruction, and axial views with 2 mm slices. A CD-ROM was distributed to each of four (4) physician/examiners (PE). All the examiners were orthopedists, specialized in spine surgery, accredited by the Sociedade Brasileira de Coluna (SBC) [Brazilian Spine Society (BSS)] and skilled in the treatment of spine fractures. Each CD-ROM contained 100 cases that were individually evaluated. There was no communication among the PEs. All the cases sent were in compliance with the inclusion and exclusion criteria mentioned above. The PEs received the original articles that describe in detail the AO/Magerl 1994,[4] TLICS,[8] and AO/2013[3] classifications, in addition to a table to be filled out individually for each case. (Attachment 1) For each patient, their clinical history, trauma mechanism, age, neurological status, and data about the integrity of the PLC were available on the CD-ROM. Concordance of the AO/2013[3] classification was performed among the 3 groups (A, B, and C) and among the eight subtypes (A1, A2, A3, A4, B1, B2, B3, and C).

The interobserver concordance for the TCLIS classification was performed based on three variables (fracture morphology, PLC injury, and neurological compromise).

For all the cases, the PEs were asked for a final decision between conservative or surgical treatment, and this information was analyzed using the Cohen's Kappa test to determine the interobserver concordance. The literature was used as a reference for orienting the values to be interpreted.[11,12] Intervals with 95% confidence were constructed for this statistic.

RESULTS

For this study, 100 cases were used, 75% of which were men and 25% women, between 20 and 60 years of age, with an average age of 35 years. The most common trauma mechanism, in 65% of the cases, was a fall from a height. The results were evaluated individually for each classification.

AO/Magerl (1994) analysis

The concordance among the PEs was evaluated taking into consideration the set of 100 imaging exams that were evaluated by all of them.

The classifications considered corresponded to the combination of type and group in the AO/Magerl 1994 classification.[4]

Based on the study results, we estimated a statistic of κ equal to 0.385, indicating marginal reproducibility. A confidence interval of 95% for the κ statistic was established by (0.363-0.407). The distribution of the results of the evaluations of the 4 PEs in the 100 cases can be seen in Figure 3.

TLICS analysis

The concordance among the examiners was evaluated considering the set of 100 imaging exams where there were evaluations from the 4 PEs.

The classification considered in this evaluation corresponded to the total score calculated using the TLICS classification, with the associated treatment options presented below:

- score from 0 to 3 - conservative.
- score equal to 4 – conservative or surgical.
- score greater than or equal to 5 - surgical.

We estimated a statistic of κ equal to 0.616 based on the results obtained in the study, indicating good reproducibility. The confidence interval of 95% for the statistic of κ was established by (0.554-0.679).

The distribution of the results can be seen in Figure 4.

The concordance among the PEs was evaluated considering the set of 100 imaging exams.

The classifications considered in this evaluation corresponded to the indication for treatment, the possible evaluation options available to each evaluator being:

– surgical treatment: yes.

– surgical treatment: no.

We estimated a statistic of κ equal to 0.690 based on the results obtained in the study, indicating good reproducibility. The confidence interval for the statistic κ was established by (0.608-0.772).

Of the 100 evaluations, the 4 PEs classified 33 cases (33.0%) as requiring conservative treatment (response of no) and 37 cases (37.0%) needing surgical treatment (answer of yes). Thus, there was concordance among all the evaluators for 70 (70%) of the imaging exams.

AO/2013 classification analysis

The concordance among the PEs was evaluated taking the set of 100 imaging exams where there were evaluations by all of them into account.

The classification considered in this evaluation corresponded to a combination of type and group. We estimated a statistic of κ equal to 0.621 based on the results obtained in the study, indicating good reproducibility. The confidence interval of 95% for the statistic κ was established by (0.583-0.659).

The distribution of results can be seen in Figure 5.

Association between the AO/2013 classification and the indication for surgery

The results relative to the percentage of surgical indications according to the AO/2013 classification can be seen in Figure 6.

DISCUSSION

Developing a classification system that is useful to all professionals who wish to better guide treatment and better understand injury mechanisms has always been the goal of many medical researchers.[3,13,14] There has always been a difficulty between using simpler systems that end up omitting some information and more complex systems that cause a lot of disagreement among professionals.

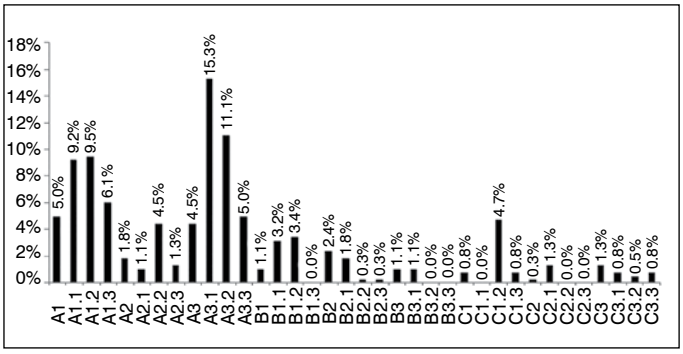

Figure 3. Percentages obtained for each subtype and group of the AO/Magerl 1994 classification.

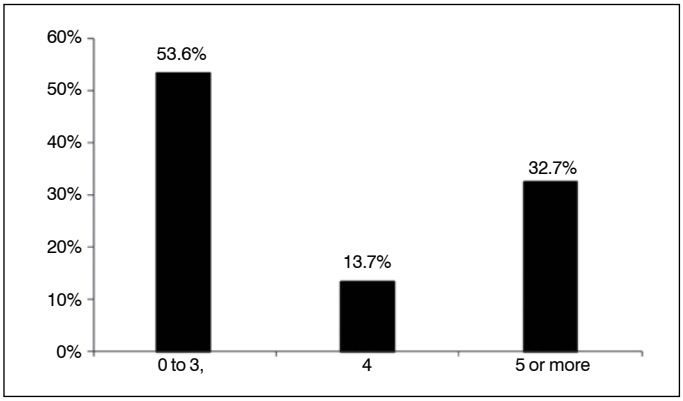

Figure 4. TLICS score percentages

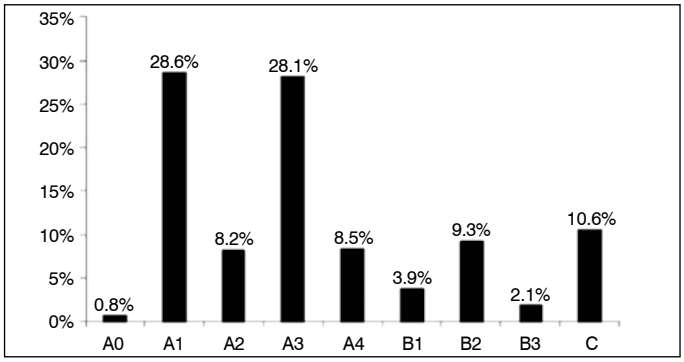

Figure 5. Percentage by type and group for the AO/2013 classification.

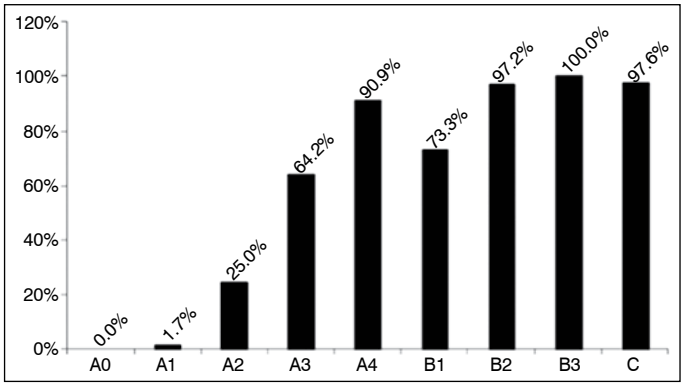

Figure 6. Surgical indication for each AO/2013 subtype.

Wood et al.[15] studied the Denis[16] and AO/Magerl 1994[4] classifications, observing only moderate reliability.

The study of thoracolumbar fractures by Blauth et al.[17] demonstrated low intra- and interobserver reliability for the AO/Magerl 1994 classification with a kappa coefficient of 0.385.

In this study, we observed a kappa coefficient of 0.385, which, although slightly higher than that from the Blauth et al.[17] study referenced above, also indicates low interobserver reliability for the AO/Magerl 1994 classification.

The AO/Magerl 1994 classification demonstrated a strong tendency towards surgical treatment for patients classified as group A3, types B and C, reserving conservative treatment for most of the A1 and A2 groups.

In the study performed with the AO/2013 classification, we obtained indications for surgical treatment in 64.7% of the type A3 fractures, more than 80% of the type B fractures, and 100% for the type C fractures, a fact that should be confronted by new studies that are using the new classification and its indication for the patient in relation to their prognosis.

The TLICS system was considered to be reliable, reproducible in smaller series, but raised questions about the cost of performing MRIs and doubts about the best treatment for a score of 4 and potential indication errors between surgical and conservative treatments.[8] The discussion about the need to perform an MRI to evaluate the integrity of the PLC arose due to the studies that proved that these injuries may go unnoticed in obese patients and those with edema.[18] Denis et al.[16] had previously related PLC injury to a worsening neurological profile and poor conservative treatment outcomes. In a recent study, the new AO/2013 classification reached a Kappa score > 0.55 in the evaluation of PLC lesions using only a clinical examination,[12] thus proving that conducting an MRI examination is not indispensable. In our study, we adopted the clinical examination as a parameter to assess PLC injury. In

our study, a TLICS score of 4 had an indication for surgery in 81.1% of the cases, mainly as a result of the clinical evaluation of the PLC injury.

Bazán et al.[14] compared the AO/Magerl 1994 and TLICS classifications and demonstrated the favorability of the latter because it is easier for physicians familiar with spine and spinal cord lesions to interpret. They used the Kappa index for statistical analysis. In our study, we obtained a k index of 0.616, which is in accordance with the references in earlier studies, indicating that the TLICS classification is more reliable than the AO/Magerl 1994 classification.

There is no consensus around the κ values that should be considered acceptable for fracture classification systems, however a value of $\kappa > 0.55$ is suggested.[13] In our study, a k value of 0.621 was found, indicating good reproducibility for the AO/2013 classification and supporting its validation.

In a similar study, Vaccaro et al.[18] found a low level of concordance for fracture morphology, almost perfect concordance for neurological compromise, and a low level of concordance for PLC lesions, resulting in a low level of concordance when considering the sum of the points. In our case series, we obtained a k coefficient of 0.616 for the final score in the TLICS classification, also indicating good reproducibility among the evaluators.

Another study demonstrated that the TLICS classification presented consensus around the definition of surgery in 47 of 49 patients (95.9%).[19] In our study, we obtained concordance of 71.1% for indications of surgery for our cases.

Joaquim et al.[19] demonstrated concordance for surgical treatment of 46.6% in their study using the TLICS classification. It was the first study to evaluate this classification using a large case series. We believe that out study had a good number of cases and is consistent with the literature.

Despite the implementation of the thoracolumbar injury severity score (TLISS), it did not prove to be statistically more reliable that the TLICS system,[11] so we did not use this score for our case series.

CONCLUSION

The new AO/2013 classification for fractures of the thoracolumbar spine was demonstrated to be a well-accepted communication tool among spine surgeons with good reproducibility, yielding a Kappa coefficient of 0.690 as compared to a k coefficient of 0.385 for AO/Magerl (1994) classification and 0.616 for TLICS. We believe that it has been established as the classification adopted by the main services that deal with this type of injury in Brazil and in other countries. However, we recognize that other studies are still necessary in different centers for it to be consolidated and for the prognosis and surgical indication for the various types of injury to be better understood.

All the authors declare that there are no potential conflicts of interest regarding this article.

CONTRIBUTIONS OF THE AUTHORS: This manuscript, which is an institutional study (Hospital do Trabalhador), has six authors. Each author made significant individual contributions the development of the manuscript. CACM and XSG were the main contributors to the writing of the manuscript. ALK, PGS, GM, and MB performed the surgeries, monitored the patients, and collected the clinical data from the medical records together with CACM. CACM and GM evaluated the data from the statistical analysis. CACM, XSG, ALK, and GM conducted the bibliographical research, the review of the manuscript, and contributed to the intellectual concept of the study.

REFERENCES

1. Zaninelli EM, Graells XSI, Néri O, Dau L. Avaliação epidemiológica das fraturas da coluna torácica e lombar de pacientes atendidos no Pronto-Socorro do Hospital do Trabalhador da UFPR de Curitiba – Paraná. Coluna/Columna. 2005;4(1):11-5.
2. Koch A, Graells XSI, Zaninelli EM. Epidemiologia de fraturas da coluna conforme seus mecanismos de trauma: análise de 502 casos Coluna/Columna. 2007;6(1):18-23.
3. Vaccaro AR, Oner C, Kepler CK, Dvorak M, Schnake K, Bellabarba C, et al. AOSpine thoracolumbar spine injury classification system: fracture description, neurological status, and key modifiers. Spine (Phila Pa 1976). 2013;38(23):2028-37.
4. 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.
5. Oner FC, Ramos LM, Simmermacher RK, Kingma PT, Diekerhof CH, Dhert WJ, et al. Classification of thoracic and lumbar spine fractures: problems of reproducibility. A study of 53 patients using CT and MRI. Eur Spine J. 2002;11(3):235-45.
6. Aebi M. Classification of thoracolumbar fractures and dislocations. Eur Spine J. 2010;19(Suppl 1):S2-7.
7. Watson-Jones R. Fractures and joint injuries. 3rd ed. London: E&S Livingstone; 1943.
8. 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.
9. Whang PG, Vaccaro AR, Poelstra KA, Patel AA, Anderson DG, Albert TJ, et al. The influence of fracture mechanism and morphology on the reliability and validity of two novel thoracolumbar injury classification systems. Spine (Phila Pa 1976). 2007;32(7):791-5.
10. Rihn JA, Yang N, Fisher C, Saravanja D, Smith H, Morrison WB, et al. Using magnetic resonance imaging to accurately assess injury to the posterior ligamentous complex of the spine: a prospective comparison of the surgeon and radiologist. J Neurosurg Spine. 2010;12(4):391-6.
11. Sethi MK, Schoenfeld AJ, Bono CM, Harris MB. The evolution of thoracolumbar injury classification systems. Spine J. 2009;9(9):780-8.
12. Urrutia J, Zamora T, Yurac R, Campos M, Palma J, Mobarec S, et al. Na independent interobserver reliability and intraobserver reproducibility evaluation of the new AOSpine Thoracolumbar Spine Injury Classification System. Spine (Phila Pa 1976). 2015;40(1):E54-8.
13. Maçaneiro CH, Miyamoto RK, Lauffer RF, Larsen RV. Avaliação da reprodutibilidade entre duas classificações de fraturas da coluna tóracolombar e suas correlações com o tratamento. Coluna/Comumna. 2008;7(2):153-9.
14. Bazán PL, Borri AE, Torres PU, Cosentino JS, Games MH. Clasificación de las fraturas toracolumbares: comparación entre las clasificaciones de AO y Vaccaro. Coluna/Columna. 2010;9(2):165-70.
15. Wood KB, Khanna G, Vaccaro AR, Arnold PM, Harris MB, Mehbod AA. Assessment of two thoracolumbar fracture classification systems as used by multiple surgeons. J Bone Joint Surg Am. 2005;87(7):1423-9.
16. 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.
17. Blauth M, Bastian L, Knop C, Lange U, Tusch G. Inter-observer reliability in the classification of thoraco-lumbar spinal injuries. Orthopade. 1999;28(8):662-81.
18. Vaccaro AR, Baron EM, Sanfilippo J, Jacoby S, Steuve J, Grossman E, et al. Reliability of a novel classification system for thoracolumbar injuries: the Thoracolumbar Injury Severity Score. Spine (Phila Pa 1976). 2006;31(11 Suppl):S62-9.
19. Joaquim AF, Fernandes YB, Cavalcante RA, Fragoso RM, Honorato DC, Patel AA. Evaluation of the thoracolumbar injury classification system in thoracic and lumbar spinal trauma. Spine (Phila Pa 1976). 2011;36(1):33-6.

Attachment 1. Table sent to participants.

Classification AO (2013)		Classification TCLIS			Classification AO (1994)		Surgical treatment	
Type	Group	Morphology	Neurological	PLC	Type	Group	Yes	No