

INTEROBSERVER REPRODUCIBILITY ASSESSMENT OF THE NEW AOSPINE CLASSIFICATION FOR SUBAXIAL CERVICAL LESIONS

AVALIAÇÃO DA REPRODUTIBILIDADE INTEROBSERVADORES DA NOVA CLASSIFICAÇÃO AOSPINE PARA LESÕES CERVICAIS SUBAXIAIS

EVALUACIÓN DE LA REPRODUCIBILIDAD INTEROBSERVADORES DE LA NUEVA CLASIFICACIÓN AOSPINE PARA LESIONES CERVICALES SUBAXIALES

LEANDRO VINÍCIUS VITAL^{1,2}, ROGÉRIO LÚCIO CHAVES DE RESENDE¹, JEFFERSON SOARES LEAL¹, RENATO DE MELO GUIMARÃES^{1,2}, ÂNGELO RIBEIRO VAZ DE FÁRIA^{1,2}

1. Universidade Federal de Minas Gerais, Hospital das Clínicas, Spine Surgery, Belo Horizonte, MG, Brazil.
2. Rede Mater Dei, Belo Horizonte, MG, Brazil.

ABSTRACT

Objective: To evaluate the interobserver agreement of the new AOSpine classification for subaxial cervical fractures. **Methods:** A descriptive study, which 11 traumatic lesions of the subaxial cervical spine (through radiographic and tomographic images), were evaluated by 16 observers being: 6 senior surgeons, 4 fellows in spinal surgery and 6 physicians residents in Orthopedics and Traumatology by the new AOSpine classification, with subsequent statistical analysis of the results. An agreement analysis was performed using the Kappa coefficient, both individually and in combination, with an interpretation of the index performed using the standardized model for Landis and Koch. To determine the level of significance of the analyzes, values less than 0.05 were considered statistically significant. **Results:** In general, the level of agreement among the examiners was considered reasonable. The lesions "A0 (F3)", "A4 (F3)", "B1", "B3", "B3 (F3)", "C", "C (F3)" and "F3" showed a low level of agreement between the examiners. The level of reasonable agreement was obtained between fractures "A0", "A1", "A4", "B2" and "C (F4)". The only fracture that presented a moderate level of agreement was the "C (F4 BL)" lesion. This result indicates that the referred injury was the fracture of the subaxial column that presented the best level of agreement among the 16 examiners in the present study. **Conclusions:** The results of the study indicate an intermediate agreement of the new AOSpine classification for subaxial cervical lesion and point to the need to carry out studies that seek to evaluate this new classification in order to better evaluate its strengths and weaknesses, contributing for its improvement. **Level of evidence III; Diagnostic study - investigation of a diagnostic test.**

Keywords: Reproducibility of Results; Fractures, bone/classification; Reliability.

RESUMO

Objetivo: Avaliar a concordância interobservadores da nova classificação AOSpine para fraturas cervicais subaxiais. **Métodos:** Estudo descritivo, por meio do qual foram avaliadas 11 lesões traumáticas da coluna cervical subaxial (através de imagens radiográficas e tomográficas) por 16 observadores, sendo seis cirurgiões seniores, quatro fellows em cirurgia da coluna e seis médicos residentes em Ortopedia e Traumatologia, pela nova classificação AOSpine. A análise de concordância foi realizada através do uso do coeficiente Kappa, tanto de forma individual como de forma combinada, sendo a interpretação do índice realizada através do modelo padronizado por Landis e Koch. Para determinar o nível de significância das análises, valores de p menores que 0.05 foram considerados estatisticamente significativos. **Resultados:** De maneira geral, o nível de concordância entre os examinadores foi considerado razoável. As lesões "A0(F3)", "A4(F3)", "B1", "B3", "B3(F3)", "C", "C(F3)" e "F3" apresentaram nível de concordância fraco entre os examinadores. O nível de concordância razoável foi obtido entre as fraturas "A0", "A1", "A4", "B2" e "C(F4)". A única fratura que apresentou nível de concordância moderado foi a lesão "C (F4 BL)". Esse resultado indica que a referida lesão foi a fratura da coluna subaxial que apresentou o melhor nível de concordância entre os 16 examinadores do presente estudo. **Conclusões:** Os resultados do estudo indicam uma concordância intermediária da nova classificação AOSpine para lesões da cervical subaxial e apontam para a necessidade da realização de estudos que busquem avaliar esta nova classificação, de forma a melhor avaliar seus pontos fortes e fracos, contribuindo para sua aprimoração. **Nível de evidência III; Estudo diagnóstico- investigação de um exame para diagnóstico.**

Descritores: Reprodutibilidade dos Testes; Fraturas Ósseas; Classificação; Confiabilidade.

RESUMEN

Objetivo: Evaluar la concordancia interobservadores de la nueva clasificación AOSpine para fracturas cervicales subaxiales. **Métodos:** Estudio descriptivo, a través del cual se evaluaron 11 lesiones traumáticas de la columna cervical subaxial (a través de imágenes radiográficas y tomográficas) por 16 observadores siendo: 6 cirujanos experimentados, 4 compañeros en cirugía de columna y 6 médicos residentes en Ortopedia y Traumatología por la nueva clasificación AOSpine. Se realizó un análisis de acuerdo utilizando el coeficiente Kappa, tanto individualmente como en combinación, con una interpretación del índice realizada utilizando el modelo estandarizado para Landis y Koch. Para determinar el nivel de significación de los análisis, se consideraron estadísticamente valores inferiores a 0,05. **Resultados:** En general, el nivel de concordancia entre los examinadores se consideró razonable. Las lesiones



"A0 (F3)", "A4 (F3)", "B1", "B3", "B3 (F3)", "C", "C (F3)" y "F3 presentaron un nivel de concordancia **débil entre** los examinadores. El nivel de concordancia razonable se obtuvo entre las fracturas A0, A1, A4, B2 y C (F4). La única fractura que presentó un nivel de concordancia moderado fue la lesión "C (F4 BL)". Este resultado indica que dicha lesión fue la fractura de la columna subaxial que presentó el mejor nivel de concordancia entre los 16 examinadores del presente estudio. Conclusiones: Los resultados del estudio indican un acuerdo intermedio de la nueva clasificación de AOSpine para lesiones cervicales subaxiales y apuntan a la necesidad de realizar estudios que busquen evaluar esta nueva clasificación, para evaluar mejor sus puntos fuertes y débil, para su mejoramiento. **Nivel de evidencia III; Estudio de diagnóstico: investigación de una prueba de diagnóstico.**

Descriptores: Reproducibilidad de los Resultados; Fractura Ósea/clasificación; Fiabilidad.

INTRODUCTION

Classifications are important in Orthopedics and Traumatology and are frequently used to guide treatment, prognosis, and communication among care teams,¹ in addition to serving to standardize the international scientific language.² To meet these objectives, the classifications must be simple, easy to apply in clinical practice, and reproducible, with high concordance among the surgeons.³

During elaboration, a classification must go through three research phases before validation for clinical use. The first phase consists of the development of the classification by a group of physicians experienced in the treatment of the particular injuries, based on several pilot studies. At this stage, concordance among the surgeons must be greater than 90%. In the next phase, the classification is tested by examiners with different levels of knowledge and practice. In the final phase it is applied clinically and documented for some period of time to confirm its validity.⁴

The most recent classification of subaxial cervical injuries is that of the AOSpine group,⁵ which seeks to be comprehensive and include various types of subaxial cervical spine bone and soft tissue injuries, based on the AOSpine classification for thoracolumbar fractures.⁶ Earlier classifications for subaxial cervical lesions, such as that of Allen and Ferguson⁷ and more recently the SLICS classification, do not present adequate interobserver concordance and leave a lot to be desired in assisting with therapeutic decisions.⁸

Bearing in mind that the subaxial cervical spine is the most common location for cervical injuries, accounting for 65% of cervical fractures and more than 75% of cervical fracture-dislocations,⁹ which are a serious public health problem and can have devastating results, such as irreversible spinal cord injury and death,⁹ a classification with good concordance among surgeons is needed to optimize the treatment of these injuries.

The objective of this study was to evaluate the interobserver concordance of the most recent subaxial injury classification (AOSpine).

METHODS

Eleven cases of traumatic subaxial cervical injury were selected from the personal files of one of the authors of the study. The associated radiographic and tomographic images contained no patient identification data. The case files were written to CDs together with a copy of the AOSpine classification (Figure 1) and delivered to the participants with a printed questionnaire, where they wrote down their responses and then returned the material to the researchers within one week.

All the participants signed an informed consent form and their identity was protected in accordance with the parameters prescribed by Resolution 466/2012 of the National Health Council / Ministry of Health, which addresses research involving human subjects. The study was approved by the Institutional Review Board of the institution.

Cases of pathological fractures (tumoral, from infection, or osteoporotic), those without information clear enough to define the classification, and injuries at multiple levels were excluded.

The cases were classified by 16 observers, 6 of whom were senior surgeons, 4 of whom were fellows in spine surgery, and 6 of whom were medical residents in Orthopedics and Traumatology.

The evaluators classified the cases according to the new AOSpine classification into types A, B, C, and F and their morphological

subtypes. (Figure 1) No clinical data were provided to the evaluators for the definition of neurological modifiers or the M modifier.

Interobserver observance was evaluated by statistical comparison of the evaluators' responses.

The analysis of concordance was conducted through use of the kappa statistic, both in individual and combined form, and the interpretation of the statistic was conducted using the interpretation model standardized by Landis and Koch (Strength of Agreement). (Figure 2)

To determine the level of significance of the analyses, values of p less than 0.05 were considered statistically significant.

All the analyses were conducted using STATA/MP version 14.0 statistical software.

The researchers were responsible for financing the research.

RESULTS

The cases presented were classified in different ways by the study examiners. (Table 1) Only one examiner (in the senior surgeon subgroup) failed to select the type of fracture in one of the study samples (case 11). The other cases were classified through the interpretation of each examiner, taking the image characteristics and the parameters indicated for each classification into account.

In general, the level of concordance among the examiners was considered fair (combined kappa = 0.2385 / p value = 0.0000) (Table 2).

Regarding the type of fracture, injuries A4(F1), B1(F3), B2(F3), C(F1,F4), C(F2), C(F2 BL), C(F2,F4), C(F3,F4), F1, F2, and F4 showed a significance level considered insignificant, with greater probability that these fractures had been classified at random (kappa < 0.00). However, these data were not statistically significant ($p > 0.05$).

Injuries A0(F3), A4(F3), B1, B3, B3(F3), C, C(F3), and F3 showed a slight level of concordance among the examiners (kappa between 0.00 and 0.20), and these values were considered significant ($p < 0.05$).

A fair level of concordance (kappa between 0.21 and 0.40) was obtained for fractures A0, A1, A4, B2, and C(F4), both with statistically significant p values ($p < 0.05$).

The only fracture that presented a moderate level of concordance was injury C(F4 BL), for which the kappa statistic value was equal to 0.4275 ($p = 0.0000$). This result indicates that this particular injury was the subaxial spine fracture with the highest level of concordance among the 16 examiners in the present study.

DISCUSSION

Classifications are useful tools, not only in the diagnosis and description of injuries, but also in defining optimal trauma team conduct, leading to better results for the individuals under care. A classification should be practical enough to be disseminated and routinely applied, but it must also be accurate, reproducible, and adequately associated with the best indication for treatment and prognosis.

In 2007, Vaccaro proposed a classification system for subaxial cervical fractures based on 3 axes: the morphology of the injury, the integrity of the posterior ligament complex, and the neurological evaluation, which generally showed good interobserver concordance, although when analyzed separately the

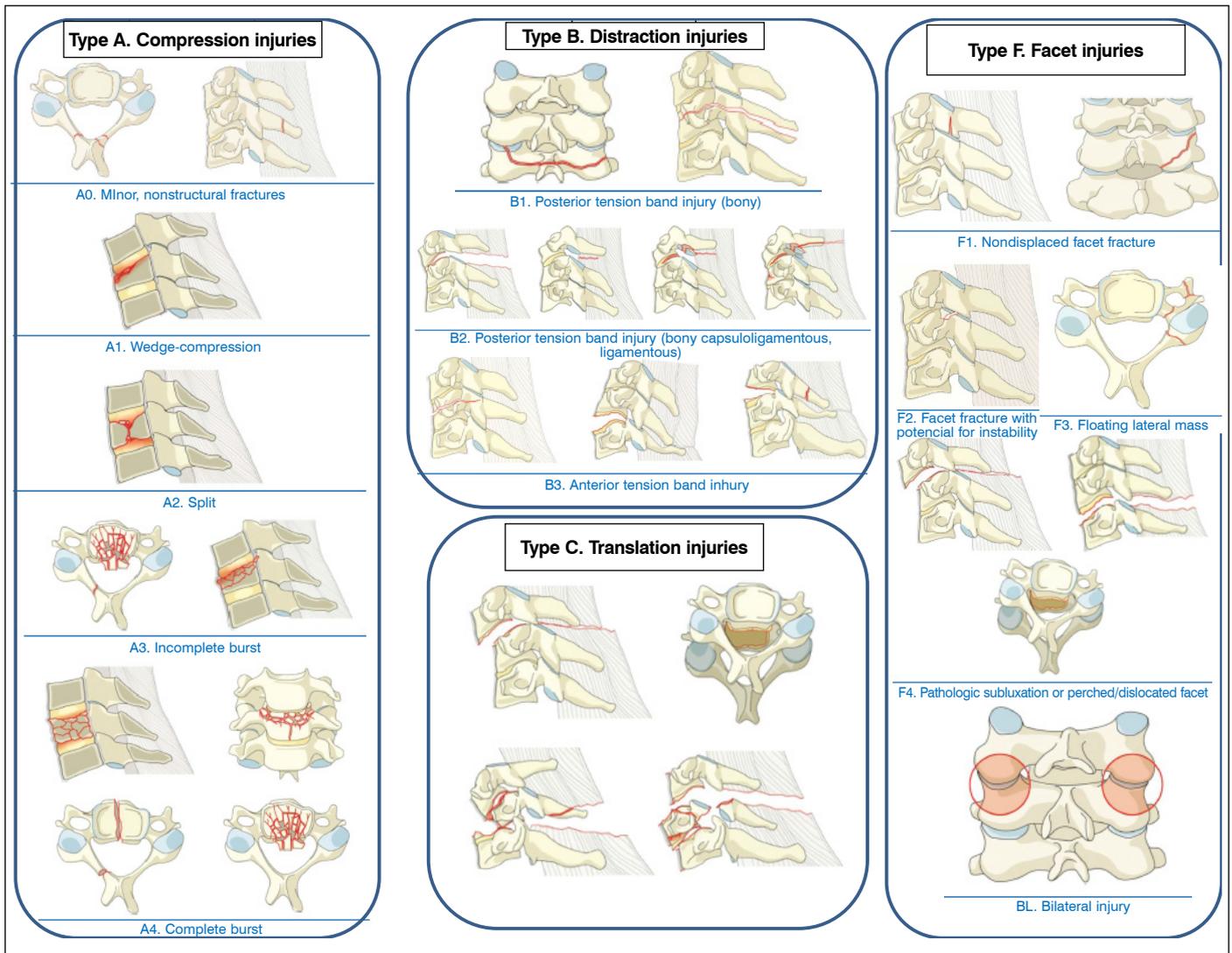


Figure 1. AOSpine classification of subaxial cervical injuries.

Kappa Statistic	Strength of agreement
<0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect

Figure 2. Landis and Koch kappa statistic interpretation model.

posterior ligament complex presented low concordance.⁸ One negative point of this classification was that it required the use of imaging examinations, which are often not readily accessible in the context of most urgent care facilities. The previously used classification was that of Allen and Ferguson (1982), who observed moderate interobserver concordance as cited by Vaccaro. In a 2008 comparative study of these two classifications, Jorge¹⁰ observed that both presented good intraobserver and moderate interobserver concordance.

Currently, the new AOSpine classification proposed by Vaccaro in 2015 has emerged as one of the most widely disseminated classifications and the results of this study offer more evidence

of its reproducibility, and, considering that the evaluations made by the physicians in this study indicate highly variable levels of concordance when comparing the different injuries included in the classification, the overall concordance was considered fair.

In 2017, Urrutia¹¹ demonstrated substantial interobserver concordance for the main fracture types (A, B, and C) and moderate concordance for the subtypes. In 2015, Vaccaro⁵ had also demonstrated good interobserver concordance for the main fracture types when describing the classification.

The data obtained show that the greater experience of the examiner does not necessarily imply higher levels of concordance, a fact that indicates that generalized use of this classification for the entire community is effective regardless of the degree of experience with these fractures.

CONCLUSIONS

From this study it was concluded that the new AOSpine classification presented a fair overall level of interobserver concordance.

The evolution of the classification is key to a more accurate approach and one with better results, thus contributing to reduce the costs involved in the process and improve the treatment of these serious injuries.

All authors declare no potential conflict of interest related to this article.

Table 1. Analysis of interobserver concordance of 11 cases of traumatic injuries of the subaxial cervical spine, as described by 16 examiners (surgeons, fellows, and residents), according to the AOSpine classification.

Type of fracture (subaxial spine)	Kappa Statistic	p value
Did not answer	-0.0057	0.5822
Minor, nonstructural fractures		
A0	0.2508	0.0000*
Minor, nonstructural fractures with Floating lateral mass		
A0(F3)	0.1183	0.0000*
Wedge-compression		
A1	0.2561	0.0000*
Complete burst		
A4	0.3751	0.0000*
Complete burst with Nondisplaced facet fracture		
A4(F1)	-0.0057	0.5822
Complete burst with Floating lateral mass		
A4(F3)	0.1814	0.0000*
Posterior tension band injury (bony)		
B1	0.1027	0.0001*
Posterior tension band injury (bony) with Floating lateral mass		
B1(F3)	-0.0057	0.5822
Posterior tension band injury (bony capsuloligamentous, ligamentous)		
B2	0.3035	0.0000*
Posterior tension band injury (bony capsuloligamentous, ligamentous) with Floating lateral mass		
B2(F3)	-0.0057	0.5822
Anterior tension band injury		
B3	0.0776	0.0024*
Anterior tension band injury with Floating lateral mass		
B3(F3)	0.0559	0.0211*
Translational injury		
C	0.1846	0.0000*
Translational injury with Nondisplaced facet fracture and Pathologic subluxation or perched/ dislocated facet		
C(F1,F4)	-0.0057	0.5822
Translational injury with Facet fracture with potential for instability		
C(F2)	-0.0173	0.7357
Translational injury with Facet fracture with potential for instability bilateral		
C(F2 BL)	-0.0057	0.5822
Translational injury with Facet fracture with potential for instability and Pathologic subluxation or perched/ dislocated facet		
C(F2,F4)	-0.0057	0.5822
Translational injury with Floating lateral mass		
C(F3)	0.2025	0.0000*
Translational injury with Floating lateral mass and Pathologic subluxation or perched/ dislocated facet		
C(F3,F4)	-0.0057	0.5822
Translational injury with Pathologic subluxation or perched/ dislocated facet		
C(F4)	0.3895	0.0000*
Translational injury with Pathologic subluxation or perched/ dislocated facet bilateral		
C(F4 BL)	0.4275	0.0000*
Nondisplaced facet fracture		
F1	-0.0115	0.6619
Facet fracture with potential for instability		
F2	-0.0057	0.5822
Floating lateral mass		
F3	0.1183	0.0000*
Pathologic subluxation or perched/ dislocated facet		
F4	-0.0057	0.5822
Combined kappa (Overall)	0.2385	0.0000*

Database = 16 examiners and 11 cases of traumatic fractures of the subaxial spine. * Represents a statistically significant difference (p<0.05), according to Fleiss' Kappa Concordance Coefficient

Table 2. Analysis of interobserver concordance of 11 cases of traumatic injuries of the subaxial cervical spine, as described by 16 examiners (surgeons, fellows, and residents), according to the AOSpine classification.

Type of fracture (subaxial spine)	Surgeons		Fellows		Residents	
	Kappa Statistic	p value	Kappa Statistic	p value	Kappa Statistic	p value
Did not answer	-0.0154	0.5783	-	-	-	-
Minor, nonstructural fractures						
A0	-0.0476	0.7296	0.6423	0.0000*	0.2548	0.0005*
Minor, nonstructural fractures with Floating lateral mass						
A0(F3)	-	-	0.3016	0.0071*	-0.0154	0.5783
Wedge-compression						
A1	-	-	-	-	0.8402	0.0000*
Complete burst						
A4	0.1750	0.0123*	-0.0233	0.5749	0.5742	0.0000*
Complete burst with Nondisplaced facet fracture						
A4(F1)	-	-	-0.0233	0.5749	-	-
Complete burst with Floating lateral mass						
A4(F3)	0.1750	0.0123*	-0.0233	0.5749	-0.0154	0.5783
Posterior tension band injury (bony)						
B1	0.1750	0.0123*	-	-	0.0419	0.2951
Posterior tension band injury (bony) with Floating lateral mass						
B1(F3)	-	-	-0.0233	0.5749	-	-
Posterior tension band injury (bony capsuloligamentous, ligamentous)						
B2	0.5594	0.0000*	0.4747	0.0001*	0.1200	0.0616
Posterior tension band injury (bony capsuloligamentous, ligamentous) with Floating lateral mass						
B2(F3)	-0.0154	0.5783	-	-	-	-
Anterior tension band injury						
B3	0.1750	0.0123*	-0.0476	0.6506	0.0921	0.1185
Anterior tension band injury with Floating lateral mass						
B3(F3)	-0.0154	0.5783	-0.0233	0.5749	-	-
Translational injury						
C	0.1200	0.0616	-0.0732	0.7239	0.4531	0.0000*
Translational injury with Nondisplaced facet fracture and Pathologic subluxation or perched/ dislocated facet						
C(F1,F4)	-0.0154	0.5783	-	-	-	-
Translational injury with Facet fracture with potential for instability						
C(F2)	-0.0312	0.6559	-	-	-0.0154	0.5783
Translational injury with Facet fracture with potential for instability bilateral						
C(F2 BL)	-0.0154	0.5783	-	-	-	-
Translational injury with Facet fracture with potential for instability and Pathologic subluxation or perched/ dislocated facet						
C(F2,F4)	-0.0154	0.5783	-	-	-	-
Translational injury with Floating lateral mass						
C(F3)	0.1370	0.0392*	0.0833	0.2492	0.4374	0.0000*
Translational injury with Floating lateral mass and Pathologic subluxation or perched/ dislocated facet						
C(F3,F4)	-0.0154	0.5783	-	-	-	-
Translational injury with Pathologic subluxation or perched/ dislocated facet						
C(F4)	0.4133	0.0000*	0.4735	0.0001*	0.0921	0.1185
Translational injury with Pathologic subluxation or perched/ dislocated facet bilateral						
C(F4 BL)	0.6743	0.0000*	0.6741	0.0000*	0.1750	0.0123*
Nondisplaced facet fracture						
F1	-0.0154	0.5783	-	-	-0.0154	0.5783
Facet fracture with potential for instability						
F2	-	-	-	-	-0.0154	0.5783
Floating lateral mass						
F3	-	-	-	-	0.3714	0.0000*
Pathologic subluxation or perched/ dislocated facet						
F4	-	-	-	-	-0.0154	0.5783
Combined kappa	0.3024	0.0000*	0.3470	0.0000*	0.3331	0.0000*

CONTRIBUTION OF THE AUTHORS: Each author made significant individual contributions to this manuscript. LVV: writing and review, surgery, statistical analysis, intellectual concept, and design of the entire research project; ARVF: data analysis and writing; RLCR: statistical analysis and review; RMG: review; JSL: writing and review.

REFERENCES

1. Drummond Filho ML, Verzani MA, Rosa AF, Pimenta CJ, Grynwald J, Cliquet Junior A. Fractures of the neck of the talus: evaluation of reproducibility of Hawkins's classification. *Acta Ortop Bras*. 2012;20(3):170-3. doi: 10.1590/S1413-78522012000300007.
2. Albuquerque RP, Giordano V, Pallottino A, Sassine T, Canedo R, Pina J, et al. Análise da reprodutibilidade das classificações das fraturas do platô tibial. *Rev Bras Ortop*. 2009;44(3):225-9. doi: 10.1590/S0102-36162009000300008.
3. Slongo T, Audige L, Schlickewei W, Clavert JM, Hunter J. Development and validation of the AO pediatric comprehensive classification of long bone fractures by the Pediatric Expert Group of the AO Foundation in collaboration with AO Clinical Investigation and Documentation and the International Association for Pediatric Traumatology. *J Pediatr Orthop*. 2006;26(1):43-9. doi: 10.1097/01.bpo.0000187989.64021.ml.
4. Audigé L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand*. 2004;75(3):184-94. doi: 10.1080/00016470412331294445.
5. Vaccaro AR, Koerner JD, Radcliff KE, Oner FC, Reinhold M, Schnake KJ, et al. AOSpine subaxial cervical spine injury classification system. *Eur Spine J*. 2016;25(7):2173-84. doi: 10.1007/s00586-015-3831-3.
6. 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.
7. Allen BL Jr, Ferguson RL, Lehmann TR, O'Brien RP. A mechanistic classification of closed, indirect fractures and dislocations of the lower cervical spine. *Spine (Phila Pa 1976)*. 1982;7(1):1-27. doi: 10.1097/00007632-198200710-00001.
8. Vaccaro AR, Hulbert RJ, Patel AA, Fisher C, Dvorak M, Lehman Jr RA, et al. The subaxial cervical spine injury classification system: a novel approach to recognize the importance of morphology, neurology, and integrity of the discoligamentous complex. *Spine (Phila Pa 1976)*. 2007;32(21):2365-74. doi: 10.1097/BRS.0b013e3181557b92.
9. Negrelli MAC, Oliveira RG, Rocha ID, Cristante AF, Marcon RM, Barros Filho TEP. Traumatic injuries of the cervical spine: current epidemiological panorama. *Acta Ortop Bras*. 2018;26(2):123-6. doi: 10.1590/1413-785220182602185460.
10. Jorge RM, Jorge GM, Karoleski MS. Reprodutibilidade da classificação de Allen versus classificação AO para fraturas da coluna cervical baixa. *Coluna/Columna*. 2008;7(4):340-3.
11. Urrutia J, Zamora T, Yurac R, Campos M, Palma J, Mobarecet S, et al. An Independent Inter- and Intraobserver Agreement Evaluation of the AOSpine Subaxial Cervical Spine Injury Classification System. *Spine (Phila Pa 1976)*. 2017;42(5):298-303. doi:10.1097/BRS.0000000000001302.