Spine

SPINE Volume 47, Number 22, pp 1541–1548

SURGERY

OPEN

Global Validation of the AO Spine Upper Cervical Injury Classification

Alexander R. Vaccaro, MD, MBA, PhD,[a] Mark J. Lambrechts, MD,[a] Brian A. Karamian, MD,[a] Jose A. Canseco, MD, PhD,[a] Cumhur Oner, MD,[b] Lorin M. Benneker, MD,[c] Richard Bransford, MD,[d] Frank Kandziora, MD, PhD,[e] Rajasekaran Shanmuganathan, MD, PhD,[f] Mohammad El-Sharkawi, MD,[g] Rishi Kanna, MD,[f] Andrei Joaquim, MD, PhD,[h] Klaus Schnake, MD,[i,j] Christopher K. Kepler, MD, MBA,[a] Gregory D. Schroeder, MD,[a] and AO Spine Upper Cervical Injury Classification International Members

Study Design. Global cross-sectional survey.

Objective. To determine the classification accuracy, interobserver reliability, and intraobserver reproducibility of the AO Spine Upper Cervical Injury Classification System based on an international group of AO Spine members.

Summary of Background Data. Previous upper cervical spine injury classifications have primarily been descriptive without

From the [a]Rothman Institute at Thomas Jefferson University, Philadelphia, PA; [b]Department of Orthopedic Surgery, University Medical Center, University of Utrecht, Utrecht, the Netherlands; [c]Sonnenhofspital Bern, University of Bern, Bern, Switzerland; [d]Department of Orthopaedic Surgery, Harborview Medical Center, University of Washington, Seattle, WA; [e]Unfallklinik Frankfurt am Main, Frankfurt, Germany; [f]Department of Orthopedics and Spine Surgery, Ganga Hospital, Coimbatore, Tamil Nadu, India; [g]Department of Orthopaedic and Trauma Surgery, Assiut University, Assiut, Egypt; [h]Department of Neurology, Neurosurgery Division, State University of Campinas, Campinas, Sao Paulo, Brazil; [i]Center for Spinal and Scoliosis Surgery, Malteser Waldkrankenhaus St. Marien, Erlangen, Germany; and [j]Department of Orthopedics and Traumatology, Paracelsus Private Medical University Nuremberg, Nuremberg, Germany.

Acknowledgment date: April 25, 2022. First revision date: May 17, 2022. Acceptance date: June 2, 2022.

This study was organized and funded by AO Spine through the AO Spine Knowledge Forum Trauma, a focused group of international Trauma experts. AO Spine is a clinical division of the AO Foundation, which is an independent medically-guided not-for-profit organization. Study support was provided directly through the AO Spine Research Department.

AO Spine Upper Cervical Injury Classification International Members: Dewan Asif, Sachin Borkar, Joseph Bakar, Slavisa Zagorac, Welege Wimalachandra, Oleksandr Garashchuk, Francisco Verdu-Lopez, Giorgio Lofrese, Pragnesh Bhatt, Oke Obadaseraye, Axel Partenheimer, Marion Riehle, Eugen Cesar Popescu, Christian Konrads, Nur Aida Faruk Senan, Adetunji Toluse, Nuno Neves, Takahiro Sunami, Bart Kuipers, Jayakumar Subbiah, Anas Dyab, Peter Loughenbury, Derek Cawley, René Schmidt, Loya Kumar, Farhan Karim, Zacharia Silk, Michele Parolin, Hisco Robijn, Al Kalbani, Ricky Rasschaert, Christian Müller, Marc Nieuwenhuijse, Selim Ayhan, Shay Menachem, Sarvdeep Dhatt, Nasser Khan, Subramaniam Haribabu, Moses Kimani, Olger Alarcon, Nnaemeka Alor, Dinesh Iyer, Michal Ziga, Konstantinos Gousias, Gisela Murray, Michel Triffaux, Sebastian Hartmann, Sung-Joo Yuh, Siegmund Lang, Kyaw Linn, Charanjit Singh Dhillon, Waeel Hamouda, Stefano Carnesecchi, Vishal Kumar, Lady Lozano Cari, Gyanendra Shah, Furuya Takeo, Federico Sartor, Fernando Gonzalez, Hitesh Dabasia, Wongthawat Liawrungrueang, Lincoln Liu, Younes El Moudni, Ratko Yurak, Héctor Aceituno, Madhivanan Karthigeyan, Andreas Demetriades, Sathish Muthu, Matti Scholz, Wael Alsammak, Komal Chandrachari, Khoh Phaik Shan, Sokol Trungu, Joost Dejaegher, Omar Marroquin, Moisa Horatiu Alexandru, Máximo-Alberto Diez-Ulloa Paulo Pereira, Claudio Bernucci, Christian Hohaus, Miltiadis Georgiopoulos, Annika Heuer, Ahmed Arieff Atan, Mark Murerwa, Richard Lindtner, Manjul Tripathi, Huynh Hieu Kim, Ahmed Hassan, Norah Foster, Amanda O'Halloran, Koroush Kabir, Mario Ganau, Daniel Cruz, Amin Henine, Jeronimo Milano, Abeid Mbarak, Arnaldo Sousa, Satyashiva Munjal, Mahmoud Alkharsawi, Muhammad Mirza, Parmenion Tsitsopoulos, Fon-Yih Tsuang, Oliver Risenbeck, Arun-Kumar Viswanadha, Samer Samy, David Orosco, Gerardo Zambito-Brondo, Nauman Chaudhry, Luis Marquez, Jacob Lepard, Juan Muñoz, Stipe Corluka, Soh Reuben, Ariel Kaen, Nishanth Ampar, Sebastien Bigdon, Damián Caba, Francisco De Miranda, Loren Lay, Ivan Marintschev, Mohammed Imran, Sandeep Mohindra, Naga Raju Reddycherla, Pedro Bazán, Abduljabbar Alhammoud, Iain Feeley, Konstantinos Margetis, Alexander Durst, Ashok Kumar Jani, Rian Souza Vieira, Felipe Santos, Joshua Karlin, Nicola Montemurro, Sergey Mlyavykh, Brian Sonkwe, Darko Perovic, Juan Lourido, Alessandro Ramieri, Eduardo Laos, Uri Hadesberg, Andrei-Stefan Iencean, Pedro Neves, Eduardo Bertolini, Naresh Kumar, Philippe Bancel, Bishnu Sharma, John Koerner, Eloy Rusafa Neto, Nima Ostadrahimi, Olga Morillo, Kumar rakesh, Andreas Morakis, Amauri Godinho, P Keerthivasan, Richard Menger, Louis Carius, Rajesh Bahadur Lakhey, Ehab Shiban, Vishal Borse, Elizabeth Boudreau, Gabriel Lacerda, Paterakis Konstantinos, Mubder Mohammed Saeed, Toivo Hasheela, Susana Núñez Pereira, Jay Reidler, Nimrod Rahamimov , Mikolaj Zimny, Devi Prakash Tokala, Hossein Elgafy , Ketan Badani, Bing Wui Ng, Cesar Sosa Juarez, Thomas Repantis, Ignacio Fernández-Bances, John Kleimeyer, Nicolas Lauper, Luis María Romero-Muñoz, Ayodeji Yusuf, Zdenek Klez, John Afolayan, Joost Rutges, Alon Grundshtein, Rafal Zaluski, Stavros I Stavridis, Takeshi Aoyama, Petr Vachata, Wiktor Urbanski, Martin Tejeda, Luis Muñiz, Susan Karanja, Antonio Martín-Benlloch, Heiller Torres, Chee-Huan Pan, Luis Duchén, Yuki Fujioka, Meric Enercan, Mauro Pluderi, Catalin Majer, Vijay Kamath.

The authors report no conflicts of interest.

Address correspondence and reprint requests to Mark J. Lambrechts, MD, Rothman Orthopaedic Institute at Thomas Jefferson University, 925 Chestnut Street, 5th Floor, Philadelphia, PA 19107; E-mail: mark.lambrechts@rothmanortho.com

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.spinejournal.com.

DOI: 10.1097/BRS.0000000000004429

incorporating a hierarchical injury progression within the classification system. Further, upper cervical spine injury classifications have focused on distinct anatomical segments within the upper cervical spine. The AO Spine Upper Cervical Injury Classification System incorporates all injuries of the upper cervical spine into a single classification system focused on a hierarchical progression from isolated bony injuries (type A) to fracture dislocations (type C).

Methods. A total of 275 AO Spine members participated in a validation aimed at classifying 25 upper cervical spine injuries through computed tomography scans according to the AO Spine Upper Cervical Classification System. The validation occurred on two separate occasions, three weeks apart. Descriptive statistics for percent agreement with the gold-standard were calculated and the Pearson χ^2 test evaluated significance between validation groups. Kappa coefficients (κ) determined the interobserver reliability and intraobserver reproducibility.

Results. The accuracy of AO Spine members to appropriately classify upper cervical spine injuries was 79.7% on assessment 1 (AS1) and 78.7% on assessment 2 (AS2). The overall intraobserver reproducibility was substantial ($\kappa = 0.70$), while the overall interobserver reliability for AS1 and AS2 was substantial ($\kappa = 0.63$ and $\kappa = 0.61$, respectively). Injury location had higher interobserver reliability (AS1. $\kappa = 0.85$ and AS2: $\kappa = 0.83$) than the injury type (AS1: $\kappa = 0.59$ and AS2: 0.57) on both assessments.

Conclusion. The global validation of the AO Spine Upper Cervical Injury Classification System demonstrated substantial interobserver agreement and intraobserver reproducibility. These results support the universal applicability of the AO Spine Upper Cervical Injury Classification System.

Key words: AO Spine, upper cervical spine, trauma, validation, reliability, reproducibility

Level of Evidence. 4

Spine 2022;47:1541–1548

Attempts at classifying upper cervical spine injuries started in 1919 when Jefferson identified potential injury mechanisms and fracture patterns of the atlas.[1] Numerous additional upper cervical spine classifications have since been proposed, but they have narrowed focus to isolated portions of the upper cervical spine.[2–9] In addition, previous injury classifications of the occipital condyles,[2,3] craniocervical junction,[4,5] atlas and transverse atlantoaxial ligament,[1,6,7] C2 peg and ring,[8–10] and C2–3 joint[11] have predominantly been descriptive with minimal ability to guide fracture management. Therefore, an upper cervical spine injury classification that is descriptive and incorporates each level of the upper cervical spine would be beneficial.

Similar to previous AO Spine classifications, the AO Spine Upper Cervical Injury Classification System follows the validation concept outlined by Audigé.[12] In short, classification systems first have experts determine the classification reproducibility and reliability. If a high reliability and reproducibility is achieved, the classification undergoes widespread international validation, which is the current step of the validation process. Subsequently, if global validation demonstrates a high degree of reliability and reproducibility, consideration then focuses on obtaining injury severity scores[13,14] to determine if the classification system can guide injury management through a treatment algorithm.[15]

Effective classification schema will result in highly accurate injury film interpretation with subsequent correct categorization of the fracture. Understanding limitations of the classification prior to global implementation is imperative in order for the classification to achieve widespread adoption. A lack of reliability and reproducibility from classification users signals the classification requires alterations prior to proceeding to the next phase of validation.[12] Therefore, the purpose of this study was to determine the international interobserver reliability and intraobserver reproducibility of the AO Spine Upper Cervical Injury Classification System.

METHODS

A Brief Description of the Classification

The AO Spine Upper Cervical Injury Classification System first divides injuries based on anatomical location. Three anatomically distinct segments are present in the upper cervical spine (I.) the occipital condyle and craniocervical junction; (II.) the C1 ring and C1–2 joint; and (III.) the C2 body, odontoid process, and C2–3 joint. Injury types are presented within each upper cervical anatomical segment. Type A injuries are predominantly bony injuries and are typically stable injury patterns. In most instances they are treated nonoperatively, but in certain circumstances they may require operative management, especially if the fracture is unlikely to heal, as is the case for dens fractures at the watershed line. Type B injuries involve a bony and/or ligamentous injury with no vertebral body translation respective to the caudal and cephalad vertebrae. These injuries are identified on computed tomography (CT) scans as a ligamentous avulsion or tension band failure. They may be stable or unstable and usually require additional imaging with dynamic radiographs or magnetic resonance imaging (MRI) to determine if operative management is indicated. Type C injuries involve either a ligamentous or bony injury that results in translation of the proximal and distal parts of the injured spinal column in any plane. These injuries are inherently unstable and frequently require operative stabilization (Fig. 1).

Classification Validation

An open call to the AO Spine community was issued to identify members willing to participate in the AO Spine Upper Cervical Injury Classification System validation. A total of 275 validation members were identified. Each participant watched a tutorial video (https://www.youtube.com/watch?v=KyUYfa_JMb4) describing the classification system and was given examples of different upper cervical spine injuries. The participants were then allowed to ask questions regarding the classification system to the supervisor (a member of the original design team of the classification system) before participating in a sample validation of three upper cervical spine injuries. Each injury was classified by the AO Spine Knowledge Forum Trauma (the

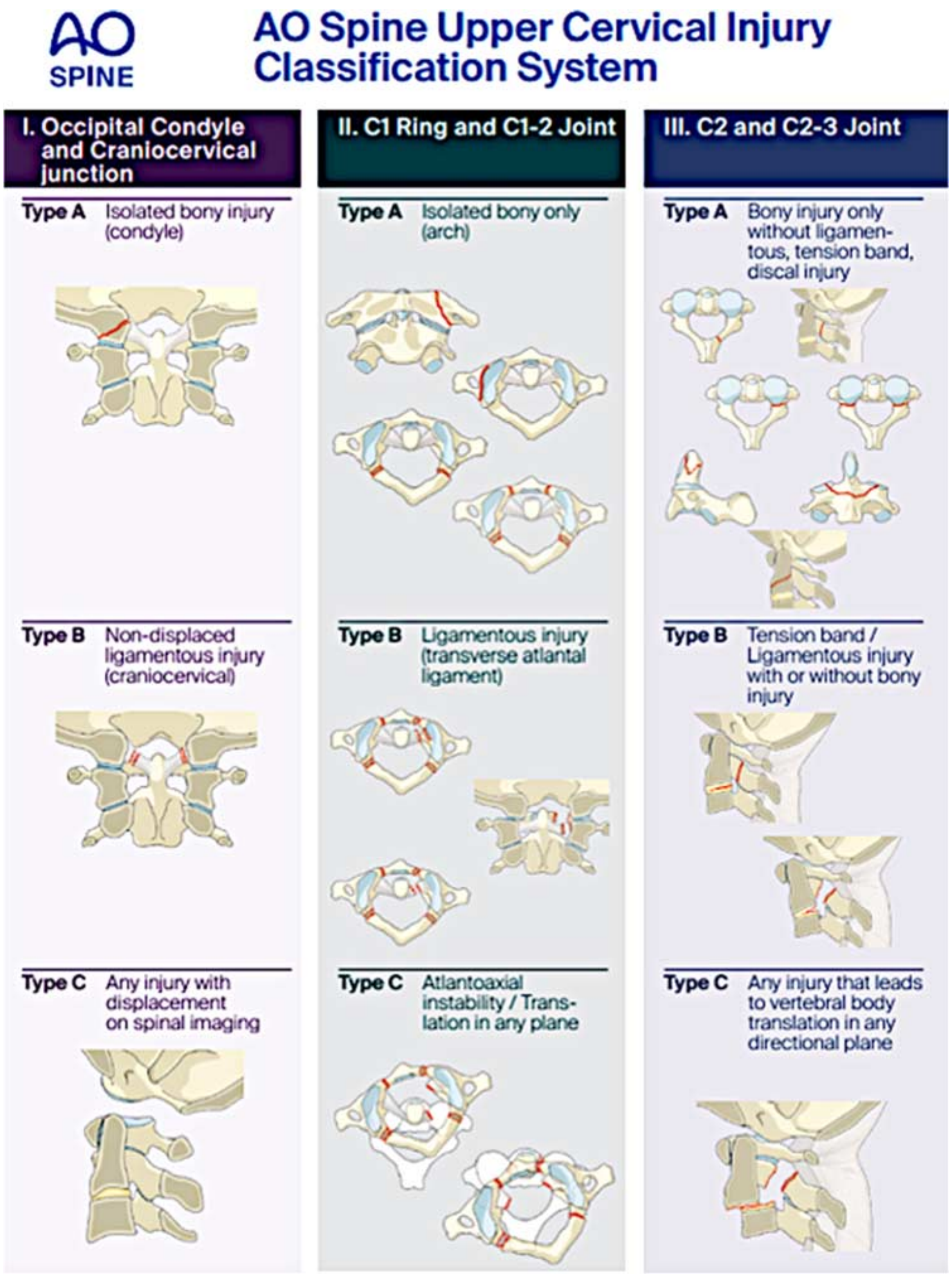

FIGURE 1. Depiction of the AO Spine Upper Cervical Injury Classification. The classification is based on injury location (occipital condyle and craniocervical junction, C1 ring and C1–2 joint, and C2 and C2–3 joint) and injury type (bony, tension band, ligamentous). Permission to use this figure was granted by the AO Foundation, AO Spine, Switzerland.

gold-standard committee) and unanimous agreement was reached prior to circulation of the injury films.

On the basis of consultation with our statistician, we attempted to provide participants with three unique injuries for each classification subtype (IA, IB, IIA, etc.) although this was not always feasible due to time constraints and an inadequate number of injury subtypes in our database. The official validation of the AO Spine Upper Cervical Injury Classification System consisted of a live, online webinar (conducted in English) with 25 unique cases showing axial, sagittal, and coronal CT videos played once at a rate of two frames/second as previously described.[16] Radiographic key images of the injuries were also provided for reference. Only injury films with a single injury were evaluated to ensure participants evaluated the correct injury, but in clinical practice if multiple injuries are present then the secondary injury should be described in parenthesis. Further, for Type B injuries, only tension band and ligamentous avulsion injuries can be evaluated with CT scan; whereas isolated ligamentous injuries without vertebral body translation require MRI or dynamic radiography for accurate classification and thus were not evaluated in this validation. An online REDCap survey captured the members' classification grades. Three weeks were allotted between the first and second assessments and the cases were re-randomized prior to the second assessment.

Statistics

Relative frequencies were tabulated based on the percent agreement between validation members and the gold-standard committee. The percent agreement was calculated for anatomic location (I, II, or III), injury type (A, B, or C), and the combination of anatomic location and injury type. Differences in relative frequencies between groups of raters were screened for potentially relevant associations with χ^2 tests in case of sufficiently large cell counts and with the Fisher exact test otherwise. Kappa coefficients (κ) were calculated based on the agreement between different validation members (interobserver reliability) and the consistency with which validation member groups chose the same classification after a three-week interval (intraobserver reproducibility). Interobserver reliability and intraobserver reproducibility were calculated for anatomical injury location, injury type, and overall classification. All of the reported kappa values utilized Fleiss' Kappa coefficient, which allows for missed ratings and comparisons between more than two validation members.[17] Interpretation of the reliability and reproducibility were based on the Landis and Koch convention, which categorized Kappa values as "slight" (< 0.2), "fair" (0.2–0.4), "moderate" (0.41–0.60), "substantial" (0.61–0.8), and "excellent" (0.81–1.0).[18]

RESULTS

After an open invitation to all AO Spine members, 275 members with varying levels of experience from each AO world region agreed to participate. A complete list of the validation members' demographics can be found in Table 1. Of the 25 cases with CT evaluations reviewed, the most commonly tested injuries were of the C1 vertebrae or C1–2 joint (N = 10) and the C2 vertebrae or C2–3 joint (N = 11), while the most common injury types were Type A (N = 10) and Type C (N = 8) (Supplemental Digital Content 1, http://links.lww.com/BRS/B896). A description of each evaluated injury film, the associated AO Upper Cervical Spine Injury Classification, and the historical injury classification are provided in Supplemental Digital Content 2, http://links.lww.com/BRS/B897.

TABLE 1. Demographics of Surgeons Who Participated in the AO Spine Upper Cervical Injury Classification

Survey Demographics	N (%)
AO region	
Number of participants	275 (100)
Africa	16 (5.8)
Asia	73 (26.5)
Central/South America	36 (13.1)
Europe	105 (38.2)
Middle East	18 (6.5)
North America	27 (9.8)
Subspecialty	
Number of participants	275 (100)
Neurosurgery	100 (36.4)
Orthopaedic spine surgery	168 (61.1)
Other	7 (2.5)
Surgical experience	
Number of participants	275 (100)
< 5 y	71 (25.8)
5–10 y	77 (28)
11–20 y	82 (29.8)
> 20 y	45 (16.4)
Work setting	
Number of participants	275 (100)
Academic	120 (43.6)
Hospital employed	120 (43.6)
Private practice	35 (12.7)
Trauma center level	
Number of participants	275 (100)
Level I	192 (69.8)
Level II	49 (17.8)
Level III	17 (6.2)
Level IV	12 (4.4)
No trauma	5 (1.8)

Gold-Standard Agreement

When assessing the agreement between validation members and the gold-standard committee, the overall classification agreement was 79.7% on assessment one (AS1) and 78.8% on assessment two (AS2). Validation members were more accurate at identifying the injury location (95.1% on AS1 and 94.1% on AS2) than the injury type (82.4% on AS1 and 82% on AS2). Although the accuracy of identifying injury location was similar regardless of anatomical location, Type B injuries (AS1: 71.2, AS2: 72.1%) were accurately identified at a much lower rate than type A (AS1: 85%, AS2: 85.7%) or type C injuries (AS1: 89.1%, AS2: 86.1%) (Table 2).

Interobserver Reliability

The overall interobserver reliability for AS1 and AS2 was substantial (κ = 0.63 and 0.61, respectively). The individual

TABLE 2. AO Spine Validation Members Percent Agreement With the Gold-Standard Committee Based on Overall Accuracy, Injury Location Accuracy, and Injury Type Accuracy

AO Spine Upper Cervical Injury Classification	Percent Agreement With Gold-Standard (%)	
	Assessment 1	Assessment 2
Overall (injury location and type)	79.7	78.8
Overall (injury location)	95.1	94.1
I	96.7	94.6
II	93.6	93.3
III	95.9	94.7
Overall (injury type)	82.4	82.0
A	85	85.7
B	71.2	72.1
C	89.1	86.1

TABLE 3. Interobserver Reliability of AO Spine Validation Members Based on Overall Classification and Injury Subtype

AO Spine Upper Cervical Injury Classification	Kappa (κ)	
	Assessment 1	Assessment 2
Overall	0.63	0.61
IA	0.75	0.70
IC	0.86	0.84
IIA	0.59	0.60
IIB	0.48	0.45
IIC	0.45	0.47
IIIA	0.69	0.67
IIIB	0.53	0.53
IIIC	0.80	0.76

injuries that had the lowest reliability were IIB (AS1: $\kappa = 0.48$ and 0.45) and IIC injuries (AS1: $\kappa = 0.45$ and 0.47). IIA (AS1: $\kappa = 0.59$ and 0.60) and IIIB injuries (AS1: $\kappa = 0.53$ and 0.53) were the only other injuries that did not reach at least substantial reliability (Table 3). After substratifying the injuries, injury location (AS1: $\kappa = 0.85$ and 0.83) had a greater interobserver reliability than injury type on AS1 and AS2. When evaluating injury type, type A (AS1: $\kappa = 0.60$; AS2: $\kappa = 0.59$) reached moderate reliability, type B had slight/moderate reliability (AS1: $\kappa = 0.41$; AS2: $\kappa = 0.39$), while type C injuries demonstrated substantial reliability (AS1: $\kappa = 0.73$; AS2: $\kappa = 0.72$) (Supplemental Digital Content 3, http://links.lww.com/BRS/B898).

Intraobserver Reproducibility

The overall intraobserver reproducibility was substantial (mean $\kappa = 0.70$). Most validation members had either excellent (38.8%) or substantial classification reproducibility (38.4%), but 15.5% had moderate reproducibility with the remainder of participants demonstrating either fair or slight reproducibility. Although 84% of validation members reached excellent intraobserver reproducibility when evaluating injury location, there was more heterogeneity for injury type. Only 33% and 35.4% of validation members reached excellent and substantial intraobserver reproducibility, respectively. An additional 22.8% of validation members demonstrated moderate reproducibility (Table 4).

DISCUSSION

Validation of the AO Spine Upper Cervical Injury Classification System demonstrated substantial interobserver reliability and intraobserver reproducibility. Nearly 80% of all injuries were correctly classified on both assessments when compared with the Gold-standard, although there was a greater accuracy at identifying injury location compared with injury type. The interobserver reliability for injury location was deemed excellent, while reliability of the injury type was moderate. Subanalysis of the injury subtypes (IA, IC, IIA, IIB, etc.) demonstrated that most injuries reached at least substantial interobserver reliability; however, all injuries to the atlas and C2 type B injuries demonstrated moderate reliability. We speculate the lower reliability for C2 type B injuries may be related to injury complexity; therefore, we discuss potential ways to distinguish Type B injuries from Type A and Type C injuries.[10]

An independent validation of the AO Spine Upper Cervical Injury Classification System was previously performed by surgeons at a single tertiary referral trauma center.[19] Similar to our results, excellent resident (range: $\kappa =$ 0.83–0.99) and attending surgeon (range: $\kappa = 0.86$–0.99) intraobserver reproducibility was identified for injury location, while injury type demonstrated substantial to excellent reproducibility for residents (range: $\kappa = 0.69$–0.92) and excellent reproducibility for attendings (range: $\kappa = 0.85$–0.98). Consistent with our results, excellent interobserver reliability was identified for injury type (range: $\kappa = 0.86$–0.88), but slightly higher interobserver reliability was demonstrated for injury type in the Maeda *et al*[19] study (AS1: $\kappa = 0.66$; AS2: $\kappa = 0.60$) compared with the results of

TABLE 4. Intraobserver Agreement for the AO Spine members' Based on the Overall Classification, Injury Location, and Injury Type

AO Spine Upper Cervical Injury Classification System*	Intraobserver Reproducibility (κ)		
	Overall Classification	Injury Location	Injury Type
Mean Kappa values (SD)	0.70 (0.19)	0.88 (0.19)	0.67 (0.22)
Level of agreement	Absolute number and percent of intraobserver agreement, N (%)		
Slight (< 0.2)	5 (2.4)	4 (1.9)	8 (3.9)
Fair (0.20–0.40)	10 (4.9)	2 (1.0)	10 (4.9)
Moderate (0.41–0.60)	32 (15.5)	10 (4.9)	47 (22.8)
Substantial (0.61–0.80)	79 (38.4)	17 (8.3)	73 (35.4)
Excellent (0.81–1.0)	80 (38.8)	173 (84.0)	68 (33)

Level of agreement is based on the Landis and Koch interpretation.
**Based on 206 validation members who participated in both assessments.*

our study (AS1: $\kappa = 0.59$; AS2: $\kappa = 0.57$). Interestingly, the results of both Maeda *et al*[19] and our study appear to indicate no "learning effect" occurs from repeat validation attempts or from additional years of surgical experience.[20] However, it should be noted the participants in the Maeda *et al*[19] study were all neurosurgeons, which may impart a benefit in classification accuracy when compared with nonspine surgeons. This was demonstrated by the ~80% classification accuracy of neurosurgeons and orthopedic spine surgeons compared with ~63% accuracy for nonspine surgeons.

Although the overall interobserver reliability and intraobserver reproducibility of the AO Spine Upper Cervical Injury Classification System was substantial, injuries to the atlas (IIA, IIB, and IIC) were identified as having lower reliability and reproducibility when compared with other injury types. Previous atlas fracture classifications have been proposed, but they have primarily been designed for descriptive purposes.[1,6] Recently, Laubach *et al*[21] found the Gehweiler classification had moderate interobserver reliability ($\kappa = 0.50$) when evaluated by 20 members of the German Society for Spine Surgeons, which was similar to the interobserver reliability obtained in our study when evaluating the AO Spine Upper Cervical Injury Classification (range: $\kappa = 0.45$–0.60 for type IIA–IIC injuries on AS1 and AS2). Therefore, it appears plausible the complexity of atlas injuries account for the moderate classification reliability regardless of the classification schema applied to these injuries.[22]

Similar to C1 ring injuries, C2 type B injuries received moderate classification reliability. These injuries have historically been labeled "atypical hangman's fractures." Unlike typical hangman's fractures, described by Levine-Edwards,[11] atypical variants are infrequently documented in the literature and have variable fracture presentation including C2 vertebral body coronal shear fractures and oblique fractures through the vertebral body, lamina, and/or pars.[23–25] These complex C2 coronal fracture variants were further described and categorized based on injury mechanisms by Effendi *et al*.[10] Multiple injury mechanisms were described (hyperextension with axial load, flexion with axial load, and flexion distraction) and they often result in AO Spine Type C injuries, based on translation of the vertebral body in either the axial or sagittal plane due to intervertebral disc injuries or avulsion fractures of the anterior or posterior longitudinal ligaments. However, the extension with axial load variant is commonly described as an atypical hangman's fracture, which is frequently classified as a Type B injury due to the tension band failure. Unfortunately, no high-quality validations of the Levine-Edwards Classification or Benzel's classification exist to compare reliability and reproducibility scores to the AO Spine Upper Cervical Injury Classification System. Similar to atlas injuries, it seems plausible classification inaccuracies of C2 injuries are due to injury complexity when compared with simple dens fractures (Type A).[24]

It is important to note that the AO Spine Upper Cervical Injury Classification System utilizes CT scans to classify all upper cervical spine injuries. This allows for minimization of the global inequality gaps in accessing MRI.[26,27] Although CT scans are quicker and more accessible than MRIs, CT scans are often limited to major trauma centers in low-income countries.[28] This may result in a persistent inability for some spine surgeons to have routine access to any advanced imaging options. In those instances, emergency departments may follow the Canadian C-Spine Rule for determining the necessity of cervical spine imaging.[29] If concerning radiographic findings are present, or if the patient is obtunded and there is concern for a cervical spine injury, patients should be transported to the nearest advanced imaging center. Although the AO Spine classification schema is based on CT evaluation, diligent use of MRI is encouraged in cases where concern for ligamentous instability exists since CT is inadequate for detecting isolated ligamentous injuries.

In particular, MRI may ultimately decide whether operative or conservative management is appropriate for Type B injuries when there is questionable injury to an intervertebral disc or ligamentous complex.

There are multiple limitations inherent to the validation of this fracture classification. First, the validation was performed by AO members, which could have inflated the overall classification accuracy, reliability and reproducibility compared with surgeons not familiar with AO classification systems. Second, the study was conducted in English and differences in fluency could have altered the validation members' ability to understand the classification system, which may have resulted in global variations in classification accuracy. Classification of the different injury types were limited to available CT scans in the AO database. Since no type IB injuries were available, they could not be evaluated by validation members which may have artificially improved the overall interobserver reliability and intraobserver reproducibility of the classification given the lower accuracy of classifying type B injuries. Finally, further attention should be given to the effect of regional variability and the influence of surgeons work setting (academic institution or level I trauma center) on the accuracy of correctly classifying injuries based on the AO Spine Upper Cervical Injury Classification System.

CONCLUSION

The international validation of the AO Spine Upper Cervical Injury Classification System demonstrated substantial interobserver reliability and intraobserver reproducibility, with excellent interobserver reliability for injury location and moderate reliability for the injury type. Although all atlas injuries demonstrated moderate interobserver reliability, this is consistent with the interobserver reliabilities of previous atlas fracture classifications. Future research targeted at understanding the reliability and reproducibility of Type IB injuries is imperative given that these injury types were not evaluated during this validation.

➢ Key Points

- ❑ The AO Spine Upper Cervical Injury Classification System has substantial intraobserver reproducibility ($\kappa = 0.70$).
- ❑ The AO Spine Upper Cervical Injury Classification System demonstrated substantial interobserver reliability on assessment one ($\kappa = 0.63$) and assessment two ($\kappa = 0.61$).
- ❑ Injury location has higher interobserver reliability on assessment one ($\kappa = 0.85$) and two ($\kappa = 0.83$) than injury type ($\kappa = 0.59$ and 0.57, respectively).
- ❑ Accurate classification of Type B injuries (71.2% accuracy on assessment one and 72.1% accuracy on assessment two) is more difficult than Type A and Type C injuries.

Acknowledgements

The authors of the manuscript would like to thank Olesja Hazenbiller for her assistance in developing the methodology and providing support during the validation. The authors would also like to thank Hans Bauer, senior biostatistician at Staburo GmbH for his assistance with the statistical analysis.

References

1. Jefferson G. Fracture of the atlas vertebra: report of four cases, and a review of those previously recorded. *Br J Surg Lond*. 1920; 7:407–22.
2. Anderson PA, Montesano PX. Morphology and treatment of occipital condyle fractures. *Spine (Phila Pa 1976)*. 1988;13:731–6.
3. Tuli S, Tator CH, Fehlings MG, Mackay M. Occipital condyle fractures. *Neurosurgery*. 1997;41:368–76.
4. Traynelis VC, Marano GD, Dunker RO, Kaufman HH. Traumatic atlanto-occipital dislocation. *Case report J Neurosurg*. 1986;65:863–70.
5. Bellabarba C, Mirza SK, West GA, et al. Diagnosis and treatment of craniocervical dislocation in a series of 17 consecutive survivors during an 8-year period. *J Neurosurg Spine*. 2006;4:429–0.
6. Gehweiler JA. The radiology of vertebral trauma. In: John A, Gehweiler Raymond L, Osborne R Jr, eds. *Frederick Becker Saunders Monographs in Clinical Radiology*, vol 16. Philadelphia, PA: Saunders; 1980:1–460.
7. Dickman CA, Greene KA, Sonntag VK. Injuries involving the transverse atlantal ligament: classification and treatment guidelines based upon experience with 39 injuries. *Neurosurgery*. 1996;38:44–50.
8. Anderson LD, D'Alonzo RT. Fractures of the odontoid process of the axis. *J Bone Joint Surg Am*. 1974;56:1663–74.
9. Benzel EC, Hart BL, Ball PA, Baldwin NG, Orrison WW, Espinosa M. Fractures of the C-2 vertebral body. *J Neurosurg*. 1994;81:206–12.
10. Effendi B, Roy D, Cornish B, Dussault RG, Laurin CA. Fractures of the ring of the axis. A classification based on the analysis of 131 cases. *J Bone Joint Surg Br*. 1981;63-B:319–27.
11. Levine AM, Edwards CC. The management of traumatic spondylolisthesis of the axis. *J Bone Joint Surg Am*. 1985;67:217–6.
12. Audigé L, Bhandari M, Hanson B, Kellam J. A concept for the validation of fracture classifications. *J Orthop Trauma*. 2005;19: 401–6.
13. Kepler CK, Vaccaro AR, Schroeder GD, et al. The Thoracolumbar AOSpine Injury Score. *Global Spine J*. 2016;6:329–4.
14. Schroeder GD, Canseco JA, Patel PD, et al. Establishing the injury severity of subaxial cervical spine trauma: validating the hierarchical nature of the AO Spine Subaxial Cervical Spine Injury Classification System. *Spine (Phila Pa 1976)*. 2021;46: 649–57.
15. Vaccaro AR, Schroeder GD, Kepler CK, et al. The surgical algorithm for the AOSpine thoracolumbar spine injury classification system. *Eur Spine J*. 2016;25:1087–94.
16. Lambrechts MJ, Schroeder GD, Karamian BA, et al. Development of Online Technique for International Validation of the AO Spine Subaxial Injury Classification System. *Global Spine J*. 2022. doi: 10.1177/21925682221098967. Epub ahead of print.
17. Gwet KL. Intrarater reliability. In: D'Agostino RB, Sullivan L, Massaro J, eds. *Wiley Encyclopedia of Clinical Trials*. Hoboken, NJ: John Wiley & Sons Inc; 2008:473–85.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
19. Maeda FL, Formentin C, de Andrade EJ, et al. Reliability of the New AOSpine Classification System for Upper Cervical Traumatic Injuries. *Neurosurgery*. 2020;86:E263–70.
20. Schroeder GD, Karamian BA, Canseco JA, et al. Validation of the AO Spine Sacral Classification System: reliability among surgeons worldwide. *J Orthop Trauma*. 2021;35:e496–501.

21. Laubach M, Pishnamaz M, Scholz M, et al. Interobserver reliability of the Gehweiler classification and treatment strategies of isolated atlas fractures: an internet-based multicenter survey among spine surgeons. *Eur J Trauma Emerg Surg*. 2020;48:601–11.
22. Fiedler N, Spiegl UJA, Jarvers JS, Josten C, Heyde CE, Osterhoff G. Epidemiology and management of atlas fractures. *Eur Spine J*. 2020;29:2477–83.
23. Starr JK, Eismont FJ. Atypical hangman's fractures. *Spine (Phila Pa 1976)*. 1993;18:1954–7.
24. Robinson AL, Möller A, Robinson Y, Olerud C. C2 fracture subtypes, incidence, and treatment allocation change with age: a retrospective Cohort study of 233 consecutive cases. *Biomed Res Int*. 2017;2017:8321680. doi:10.1155/2017/8321680.
25. Al-Mahfoudh R, Beagrie C, Woolley E, et al. Management of typical and atypical Hangman's fractures. *Global Spine J*. 2016;6:248–56.
26. Ogbole GI, Adeyomoye AO, Badu-Peprah A, Mensah Y, Nzeh DA. Survey of magnetic resonance imaging availability in West Africa. *Pan Afr Med J*. 2018;30:240.
27. Volpi G. Radiography of diagnostic imaging in Latin America. *Nucl Med Biomed Imaging*. 2016;1:10–2.
28. Hricak H, Abdel-Wahab M, Atun R, et al. Medical imaging and nuclear medicine: a Lancet Oncology Commission. *Lancet Oncol*. 2021;22:e136–72.
29. Stiell IG, Wells GA, Vandemheen KL, et al. The Canadian C-spine rule for radiography in alert and stable trauma patients. *JAMA*. 2001;286:1841–8.