







The Influence of Regional Differences on the Reliability of the AO Spine Sacral Injury Classification System

Global Spine Journal
2022, Vol. 0(0) 1–8
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/21925682211068419
journals.sagepub.com/home/gsj


Brian A. Karamian, MD¹ , Gregory D. Schroeder, MD¹, Mark J. Lambrechts, MD¹ , Jose A. Canseco, MD, PhD¹  AO Spine Sacral Classification Group Members, Emiliano N. Vialle, MD² , Shanmuganathan Rajasekaran, PhD³ , Lorin M. Benneker, MD⁴, Marcel R. Dvorak, MD, FRCSC⁵, Frank Kandziora, MD⁶, Cumhur Oner, MD⁷, Klaus Schnake, MD⁸, Christopher K. Kepler, MD, MBA¹, and Alexander R. Vaccaro, MD, PhD, MBA¹

Abstract

Study Design: Global cross-sectional survey.

Objective: To explore the influence of geographic region on the AO Spine Sacral Classification System.

Methods: A total of 158 AO Spine and AO Trauma members from 6 AO world regions (Africa, Asia, Europe, Latin and South America, Middle East, and North America) participated in a live webinar to assess the reliability, reproducibility, and accuracy of classifying sacral fractures using the AO Spine Sacral Classification System. This evaluation was performed with 26 cases presented in randomized order on 2 occasions 3 weeks apart.

Results: A total of 8320 case assessments were performed. All regions demonstrated excellent intraobserver reproducibility for fracture morphology. Respondents from Europe ($k = .80$) and North America ($k = .86$) achieved excellent reproducibility for fracture subtype while respondents from all other regions displayed substantial reproducibility. All regions demonstrated at minimum substantial interobserver reliability for fracture morphology and subtype. Each region demonstrated >90% accuracy in classifying fracture morphology and >80% accuracy in fracture subtype compared to the gold standard. Type C morphology ($p_2 = .0000$) and A3 ($p_1 = .0280$), B2 ($p_1 = .0015$), C0 ($p_1 = .0085$), and C2 ($p_1 = .0016$, $p_2 = .0000$) subtypes showed significant regional disparity in classification accuracy ($p_1 =$ Assessment 1, $p_2 =$ Assessment 2). Respondents from Asia (except in A3) and the combined group of North, Latin, and South America had accuracy percentages below the combined mean, whereas respondents from Europe consistently scored above the mean.

Conclusions: In a global validation study of the AO Spine Sacral Classification System, substantial reliability of both fracture morphology and subtype classification was found across all geographic regions.

Keywords

sacrum, classification, AO spine, international, geographic region, AO Trauma, reliability, injury morphology, injury subtype, intraobserver agreement

¹Rothman Institute at Thomas Jefferson University, Philadelphia, PA, USA

²Spine Surgery Group, Department of Orthopaedics, Cajuru University Hospital, Catholic University of Parana, Curitiba, Brazil

³Department of Orthopedics and Spine Surgery, Ganga Hospital, Coimbatore, India

⁴Department of Orthopaedic Surgery, Inselspital, University of Bern, Bern, Switzerland

⁵Division of Spine, University of British Columbia

⁶Unfallklinik Frankfurt am Main, Frankfurt, Germany

⁷Department of Orthopedic Surgery, University Medical Center, University of Utrecht, Utrecht, Netherlands

⁸Center for Spinal Surgery, Schön Klinik Nürnberg Fürth, Fürth, Germany

Corresponding Author:

Mark J. Lambrechts, MD, Rothman Orthopaedic Institute at Thomas Jefferson University 925 Chestnut St, 5th Floor, Philadelphia, PA 19107, USA.

Email: mark.lambrechts@rothmanortho.com



Creative Commons Non Commercial No Derivs CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

The sacrum serves as the keystone between the spine and pelvis, transmitting axial loads from the spinal column to the ilia. As a result, approximately 10–45% of patients with traumatic pelvic injuries have concomitant sacral fractures.¹ While relatively rare injuries, the detection of sacral fractures has improved considerably as a result of advances in general trauma management and diagnostic imaging modalities, increasing the published incidence of sacral fractures.² Nevertheless, the management of these injuries remains controversial. No high-level studies have investigated the relationship between sacral fracture management and long-term clinical outcomes. As a result, few scientifically based insights on sacral fracture management can be gathered from the current literature. This, in part, is due to the lack of a comprehensive classification system to facilitate in the communication, research, and education of sacral fractures. While numerous sacral classification systems have been described, the clinically relevant elements of sacral fractures have failed to be captured by previous schemes.^{3–9}

The AO Spine Knowledge Forum Trauma in conjunction with a group of orthopedic pelvic trauma surgeons from AO Trauma recently developed the AO Spine Sacral Classification System, a concise and comprehensive classification scheme to facilitate the standardization of treatment and prognostication of outcomes after sacral injury.¹⁰ The classification system is separated into 3 morphologic fracture categories: type A (lower sacrococcygeal injuries), type B (posterior pelvic injuries), and type C (spino-pelvic injuries). Similar to previous AO Spine classification systems, it is designed in a hierarchical manner in which each morphologic type is subdivided into increasing numerical subtypes based on the severity of injury.^{11,12} Case-specific modifiers and neurologic injury at the time of examination incorporate patient-specific data to individualize management within a universally applicable scheme [Figure 1](#).

Many factors may play a role in the ability to diagnose and classify sacral fractures. Cultural expectations, available resources, and local provider education and training may all affect injury interpretation and are intrinsically tied to geographic region. In order to validate the AO Spine Sacral Classification system as a universal tool to serve as the basis for the development of globally applicable treatment algorithms, the regional differences in the radiographic interpretation of sacral fractures must be elucidated. Accordingly, the aim of this study is to explore the potential influence of geographic region on the classification of sacral fractures using the AO Spine Sacral Classification System.

Methods

A survey was sent to 172 AO Spine and AO Trauma members from all AO Spine world regions (Europe, North America, South America, Africa, Asia [inclusive of Asia-Pacific], and the Middle East) who routinely treat sacral fractures. For training

purposes, the respondents were provided educational resources on each component of the classification system, which included a video introduction, written and pictorial classification descriptions, and a 10-case practice assessment. In an effort to limit regional bias, all injury patterns were described thoroughly using both visual and verbal modalities. Previously unclassified images were provided by AO Spine from a DICOM database. These injury films were reviewed and classified by members of the AO Knowledge Forum Trauma. Only injuries with complete agreement from the AO Knowledge Forum Trauma were deemed acceptable for validation use and their injury classification was deemed the “gold standard.” All remaining cases with incomplete agreement or poor-quality imaging were excluded. Respondents participated in a live webinar assessing high-resolution key images as well as axial/sagittal/coronal computed tomography (CT) videos for 26 sacral fractures. This evaluation was performed on 2 separate occasions 3 weeks apart, where identical cases were presented in randomized order. At minimum, 2 cases of each fracture subtype were included. The case order was randomized in both assessments such that a consecutive series was not presented given the hierarchical nature of the classification system.

Statistical Analysis

The summary of surgeon respondent demographics including years of practice, cases per year, and surgical specialty were analyzed using chi-squared tests. Cohen’s kappa (k) statistic was used to assess the reliability of classification between independent observers (interobserver agreement) and reproducibility between classifications of the same observer for separate evaluations (intraobserver reproducibility). Interobserver agreement and intraobserver reproducibility were calculated for each injury morphology (A, B, and C) and injury subtype (A1, A2, A3, B1, B2, B3, C0, C1, C2, and C3) and stratified by region. The k coefficients were interpreted using the Landis and Koch grading system with $k < .20$ defined as slight reliability/reproducibility, $.20–.40$ as fair reliability/reproducibility, $.40–.60$ as moderate reliability/reproducibility, $.60–.80$ as substantial reliability/reproducibility, and $>.80$ as excellent reliability/reproducibility.¹³

Accuracy of classification was calculated through percentage agreement with the “gold standard” fracture type (morphology and subtype) for each assessment. Gold standard agreement was stratified by region and compared with a Fischer Exact test. The regions including the Middle East and Africa, as well as North America and Latin/South America, were combined for the analysis of gold standard agreement due to the low number of participants in these regions. Statistical significance was defined at $P < .05$ and no adjustment for multiplicity was performed.

Results

Of the 172 surgeons invited to participate, 158 completed the first assessment and 162 completed the second assessment for

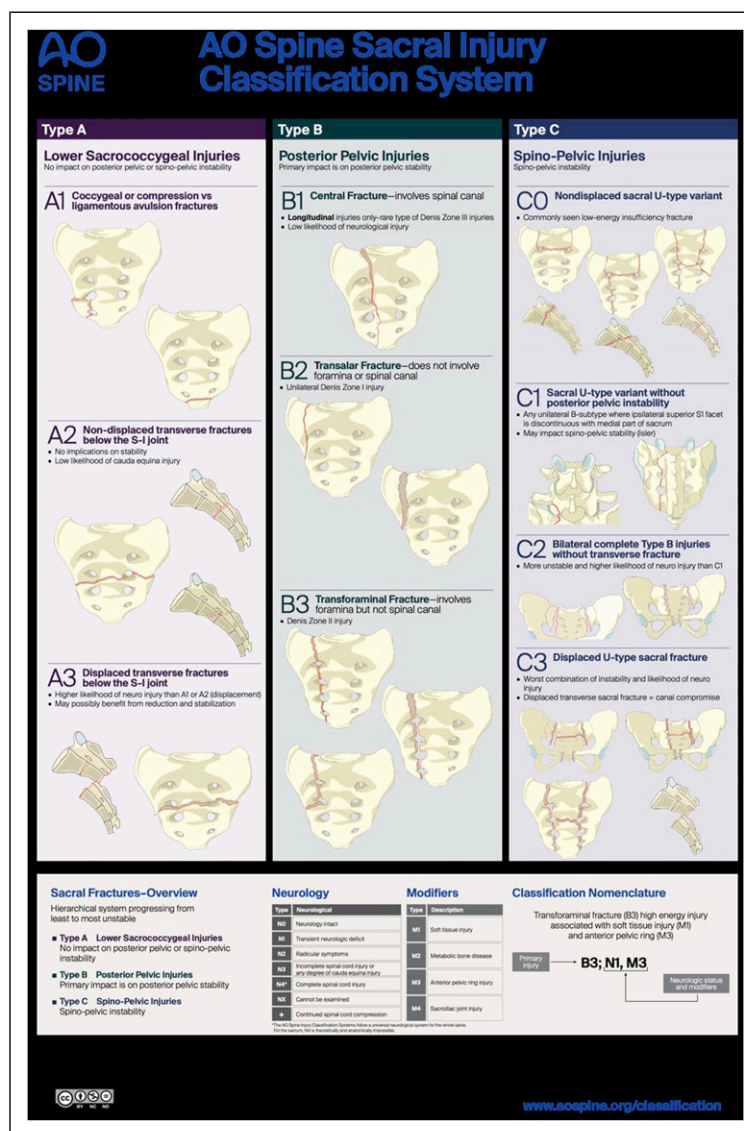


Figure 1. AO Spine sacral injury classification poster describes the hierarchal nature of sacral fractures and the resultant stability they possess. Permission to use this image was granted by the AO Foundation©, AO Spine, Switzerland.

a total of 8320 case assessments. The largest portion of respondents practiced in Europe (42.4%), followed by Asia (20.3%) and Latin and South America (19.2%), with smaller contributions from the Middle East (8.1%), Africa (5.25%), and North America (4.7%). Regardless of world region, most respondents had been in practice for 11–20 years, with the least amount of respondents having more than 20 years of experience. Additionally, respondents from the Middle East were more likely to have either 11–20 or >20 years of experience, North America respondents were more likely to have been in practice for <5 years, while the remaining regions had a more even distribution of years in practice ($P = .026$). The majority of respondents practiced in an academic or hospital-employed setting, except for Latin and South America, which predominantly were employed in the private practice settings ($P < .001$). Of all respondents, the majority were orthopedic

spine surgeons, followed by neurosurgeons, with a small subset consisting of general surgeons and trauma surgeons ($P = .001$) [Table 1](#). The 26 cases for review consisted of 7 (26.9%) type A, 8 (30.8%) type B, and 11 (42.3%) type C fractures with 2 cases each represented for A1, A2, B1, C0 subtypes and 3 cases each represented for A3, B2, B3, C1, C2, and C3 subtypes.

Intraobserver Reproducibility

Surgeons from all 6 geographic regions demonstrated excellent intraobserver reproducibility for sacral fracture morphology classification [Table 2](#). When evaluating fracture subtype, respondents from North America ($k = .86$) achieved excellent intraobserver reproducibility while respondents from all other regions displayed substantial reproducibility

Table 1. Summary of Surgeon Respondent Demographics.

| | Africa | Asia | Europe | Latin/South America | Middle East | North America | Total Respondents | P- value |
|--------------------------|------------|------------|------------|---------------------|-------------|---------------|-------------------|--------------------|
| Years of Practice | | | | | | | | .026 ^a |
| <5 | 6 (8.1%) | 20 (27.0%) | 18 (24.3%) | 17 (23.0%) | 5 (6.8%) | 8 (10.8%) | 74 | |
| 5–10 | 3 (3.8%) | 28 (35.4%) | 28 (35.4%) | 16 (20.3%) | 1 (1.3%) | 3 (3.8%) | 79 | |
| 11–20 | 6 (6.5%) | 25 (27.2%) | 29 (31.5%) | 22 (23.9%) | 8 (8.7%) | 2 (2.2%) | 92 | |
| >20 | 4 (6.6%) | 14 (23.0%) | 22 (36.1%) | 11 (18.0%) | 10 (16.4%) | 0 (.0%) | 61 | |
| Cases per year | | | | | | | | .277 |
| 0–5 | 7 (5.6%) | 39 (31.0%) | 38 (30.2%) | 29 (23.0%) | 7 (5.6%) | 6 (4.8%) | 126 | |
| 6–10 | 10 (10.3%) | 23 (23.7%) | 30 (30.9%) | 25 (25.8%) | 5 (5.2%) | 4 (4.1%) | 97 | |
| 11–20 | 1 (1.9%) | 16 (29.6%) | 19 (35.2%) | 8 (14.8%) | 9 (16.7%) | 1 (1.9%) | 54 | |
| >20 | 1 (3.4%) | 9 (31.0%) | 10 (34.5%) | 4 (13.8%) | 3 (10.3%) | 2 (6.9%) | 29 | |
| Specialty | | | | | | | | .001 ^a |
| General | 2 (10.5%) | 6 (31.6%) | 8 (42.1%) | 3 (15.8%) | 0 (.0%) | 0 (.0%) | 19 | |
| Neurosurgery | 2 (2.9%) | 9 (13.2%) | 18 (26.5%) | 26 (38.2%) | 9 (13.2%) | 4 (5.9%) | 68 | |
| Ortho spine | 14 (7.1%) | 65 (32.8%) | 62 (31.3%) | 37 (18.7%) | 12 (6.1%) | 8 (4.0%) | 198 | |
| Trauma | 1 (5.0%) | 7 (35.0%) | 7 (35.0%) | 0 (.0%) | 3 (15.0%) | 0 (.0%) | 18 | |
| Other | 0 (.0%) | 1 (50.0%) | 1 (50.0%) | 0 (.0%) | 0 (.0%) | 1 (50.0%) | 3 | |
| Practice setting | | | | | | | | <.001 ^a |
| Private | 3 (6.3%) | 11 (22.9%) | 4 (8.3%) | 27 (56.3%) | 2 (4.2%) | 1 (2.1%) | 48 | |
| Academic | 8 (6.8%) | 40 (33.9%) | 35 (29.7%) | 11 (9.3%) | 12 (10.2%) | 12 (10.2%) | 128 | |
| Hospital employed | 8 (5.7%) | 37 (26.2%) | 58 (41.1%) | 28 (19.9%) | 10 (7.1%) | 0 (.0%) | 141 | |

^aIndicates statistical significance (P < .05).

Table 2. Intraobserver Reproducibility Mean Kappa Values by Geographic Region.

| Region | Africa n = 9 | Asia n = 35 | Europe n = 73 | Latin and South America n = 33 | Middle East n = 14 | North America n = 8 | All Regions n = 172 |
|----------------------------|-----------------|----------------|------------------|-----------------------------------|-----------------------|------------------------|------------------------|
| Fracture morphology | .89 | .86 | .90 | .81 | .89 | .96 | .87 |
| Fracture subtype | .68 | .75 | .80 | .71 | .78 | .86 | .77 |

Table 2. Respondents from Latin and South America demonstrated the lowest reproducibility in classifying fracture morphology (k = .81) and respondents from Africa exhibited the lowest reproducibility in classifying fracture subtype (k = .68).

Interobserver Reliability

Overall, excellent reliability was found in the classification of fracture morphology (k = .842/.850, Assessment 1/ Assessment 2) and substantial reliability was found for fracture subtype (k = .719/.751) when all regions were evaluated for interobserver reliability. A slight improvement was also noted between the fracture morphology reliability on assessment 2 (k = .850) compared to assessment 1 (k = .842) and this was also seen when evaluating fracture subtype: assessment 1 (k = .719) and assessment 2 (k = .751). Respondents from Europe (k = .859/.901), the Middle East (k = .868/.873), and North America (k = .984/.889) demonstrated excellent interobserver reliability in classification of all fracture morphologies for both assessments [Table 3](#). Surgeons

practicing in Asia had excellent interobserver reliability (k = .821/.819) for the combined morphology classification. Respondents from Africa (k = .825/.767) and Latin and South America (k = .785/.780) achieved at minimum substantial overall interobserver reliability for the classification of morphology. When classifying fracture subtype, all regions attained at minimum substantial reliability in both evaluations. Raters from North America demonstrated the greatest reliability (k = .881/.801) while raters from Africa and Latin and South American demonstrated the lowest reliability (k = .650/.665 and k = .660/.671, respectively) [Table 3](#).

Gold Standard Agreement

Each region correctly classified morphology in greater than 90% of cases and fracture subtype in over 80% of cases in both assessments according to the gold standard [Table 4](#). However, accuracy in classifying fracture morphology and fracture subtype varied significantly between regions in the second assessment (p₂ = .0000; p₁ = Assessment 1,

Table 3. Interobserver Agreement Analysis by Geographic Region.

| Region Assessment Number | Africa | | Asia | | Europe | | Latin and South America | | Middle East | | North America | | All Regions | |
|--------------------------|----------------------------|------|------|------|--------|------|-------------------------|------|-------------|------|---------------|------|-------------|------|
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| | Fracture morphology | | | | | | | | | | | | | |
| A | .957 | .793 | .936 | .920 | .955 | .964 | .934 | .889 | .943 | .925 | 1.000 | .917 | .948 | .927 |
| B | .737 | .681 | .745 | .754 | .800 | .859 | .696 | .713 | .819 | .823 | .974 | .857 | .775 | .794 |
| C | .782 | .828 | .781 | .782 | .822 | .879 | .725 | .739 | .842 | .871 | .977 | .894 | .802 | .828 |
| Overall/Combined | .825 | .767 | .821 | .819 | .859 | .901 | .785 | .780 | .868 | .873 | .984 | .889 | .842 | .850 |
| Fracture subtype | | | | | | | | | | | | | | |
| A1 | 1.000 | .870 | .874 | .855 | .892 | .955 | .850 | .822 | .802 | .851 | 1.000 | .927 | .881 | .894 |
| A2 | .708 | .585 | .727 | .672 | .759 | .756 | .725 | .691 | .736 | .789 | .917 | .671 | .743 | .719 |
| A3 | .778 | .512 | .781 | .611 | .800 | .741 | .751 | .745 | .588 | .787 | .948 | .784 | .773 | .713 |
| B1 | .768 | .657 | .543 | .650 | .696 | .689 | .670 | .534 | .706 | .435 | .927 | .782 | .678 | .634 |
| B2 | .581 | .588 | .759 | .754 | .775 | .917 | .762 | .701 | .830 | .801 | 1.000 | .831 | .779 | .815 |
| B3 | .581 | .713 | .829 | .846 | .804 | .870 | .607 | .703 | .831 | .942 | 1.000 | .900 | .771 | .826 |
| C0 | .526 | .637 | .589 | .653 | .639 | .738 | .473 | .577 | .659 | .847 | .630 | .632 | .598 | .681 |
| C1 | .484 | .679 | .547 | .607 | .608 | .728 | .479 | .538 | .686 | .723 | .835 | .712 | .584 | .663 |
| C2 | .570 | .746 | .687 | .780 | .867 | .927 | .696 | .669 | .856 | .934 | .844 | .892 | .791 | .836 |
| C3 | .508 | .664 | .575 | .630 | .599 | .756 | .585 | .732 | .541 | .799 | .711 | .877 | .593 | .729 |
| Overall/Combined | .650 | .665 | .691 | .706 | .744 | .808 | .660 | .671 | .723 | .791 | .881 | .801 | .719 | .751 |

Table 4. Comparison to Gold Standard Fracture Classification by Geographic Region.

| Region Assessment Number | Asia | | Europe | | Middle East and Africa | | North America, Latin, and South America | | P-value | | All Regions | |
|--------------------------|----------------------------|-------|--------|-------|------------------------|-------|---|-------|---------|-------|-------------|-------|
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| | Fracture morphology | | | | | | | | | | | |
| A | 97.7% | 97.7% | 98.7% | 98.8% | 99.2% | 97.2% | 98% | 98.2% | .6217 | .4844 | 98.4% | 98.2% |
| B | 90.9% | 97.7% | 94.3% | 98.8% | 93.5% | 97.2% | 93.2% | 98.2% | .4017 | .4844 | 93.3% | 98.2% |
| C | 91% | 91.2% | 91.8% | 96.4% | 92.6% | 94.1% | 89% | 89.9% | .3684 | .0000 | 91.1% | 93.5% |
| Overall/Combined | 93.2% | 93.4% | 94.9% | 96.3% | 95.1% | 93.5% | 93.4% | 92.6% | .1776 | .0000 | 94.2% | 94.5% |
| Fracture subtype | | | | | | | | | | | | |
| A1 | 92.1% | 95.2% | 94.9% | 98.6% | 94.3% | 97.6% | 90.4% | 97.4% | .6193 | .4961 | 93.2% | 97.5% |
| A2 | 90.2% | 87.1% | 90.4% | 92.9% | 91.7% | 92.5% | 91.8% | 92.2% | .9819 | .5969 | 90.8% | 91.5% |
| A3 | 86.7% | 69.9% | 88.7% | 78.5% | 73.1% | 79.4% | 88.2% | 80.3% | .0280 | .2782 | 86.4% | 77.4% |
| B1 | 71.7% | 87.5% | 86% | 85% | 83.8% | 77.5% | 87.5% | 81.8% | .0582 | .5347 | 83.3% | 83.8% |
| B2 | 90.4% | 85.4% | 90.6% | 95.2% | 92% | 81.7% | 93.7% | 84.2% | .8006 | .0015 | 91.5% | 89% |
| B3 | 96.6% | 92.3% | 94.1% | 91.3% | 90.2% | 93.1% | 88.3% | 86.1% | .0944 | .3617 | 92.7% | 90.5% |
| C0 | 78.3% | 92.1% | 80% | 97.9% | 76.5% | 97.5% | 72.2% | 87.2% | .6431 | .0085 | 77.4% | 94.1% |
| C1 | 73.9% | 79.3% | 73.2% | 85.7% | 80.8% | 85.2% | 70% | 73.3% | .5486 | .0368 | 73.4% | 81.4% |
| C2 | 79.1% | 82.6% | 94.1% | 96.2% | 88.2% | 88.5% | 84.4% | 79.5% | .0016 | .0000 | 88.1% | 88.5% |
| C3 | 57.6% | 66.3% | 57.6% | 75.7% | 51% | 80% | 66.1% | 81.7% | .2792 | .0621 | 58.9% | 75.9% |
| Overall/Combined | 81.7% | 83.8% | 84.9% | 89.7% | 82.1% | 87.3% | 83.3% | 84.4% | .1956 | .0000 | 83.6% | 87% |

p_2 = Assessment 2). Specifically, Type C morphology (p_2 = .0000) and subtypes A3 (p_1 = .0280, p_1 = Assessment 1), B2 (p_1 = .0015), C0 (p_1 = .0085), and C2 (p_1 = .0016, p_2 = .0000) showed significant regional disparity in the percentage classified correctly. In fracture subtypes with significant differences in classification accuracy, respondents from Asia (except in A3) and the combined group of North, Latin, and South America consistently had accuracy percentages below

the combined mean, whereas respondents from Europe always scored above the combined mean [Table 4](#).

Discussion

Management of spinal injuries has been shown to vary between global regions.¹⁴⁻¹⁷ Differences in injury interpretation, perceived severity, cultural expectations, available resources,

costs of surgery, and local provider education may all play a role in the regional disparity present in spinal trauma management. Previous literature has touched upon the inconsistencies in spine trauma management between modern and emerging countries.^{16,18,19} In an effort to understand the causes of such disparity, a better understanding of the differences in regional interpretation of spinal injury and trauma is first required. Accordingly, the goal of this study is to explore the potential influence of geographic region on the reliability and reproducibility of sacral fracture classification using the AO Spine Sacral Classification System. Surveying surgeons from around the world, our validation results demonstrate substantial reliability of both fracture morphology and fracture subtype despite variations in geography. These results are essential in laying the foundation for the development of globally accepted treatment algorithms to ultimately accompany the AO Spine Sacral Classification System. Regional differences in the care and treatment of spinal trauma, including indications for operative management, must be taken into consideration during the development of a classification scheme. The results of this study underscore the universal nature of the AO Spine Sacral Classification System, without the need for further refinement based upon regional interpretation.

Despite the overall substantial reliability of the classification system, subtle differences in the performance between world regions were noted. When evaluating the interobserver reliability of fracture classification by region, observers from Africa and Latin and South America demonstrated less than excellent reliability on 1 or both assessments for fracture morphology. Similarly, these regions also demonstrated the lowest reliability scores, albeit substantial, for classification of fracture subtype as well as reproducibility of fracture subtype. Furthermore, observers from North, Latin, and South America, as well as Asia, had significantly more incorrect classifications compared to all other regions when evaluating accuracy of classification in comparison to the gold standard. Contiguous regions of the Americas were combined for statistical analysis given the small number of participating observers from North America. Given that North America was the only region with excellent reliability for both fracture morphology and sub-type, it suggests that the accuracy of Latin and South American observers may be less than the combined performance presented. Moreover, surgeons from Africa exhibited the largest difference in reproducibility between fracture morphology ($k = .89$) and subtype ($k = .68$) which highlights the difficulty in interpreting the distinctive patterns of sacral fractures required for stratification.

As this is the first study to evaluate the global interpretation of sacral fractures, only suggestions as to the reasons for the aforementioned differences in regional interpretation can be proposed. One reason for differences in reliability, reproducibility, and accuracy may be due to worldwide variations in imaging infrastructure, with decreased availability of advanced imaging modalities in rural areas of middle- and

low-income countries.^{20,21} Given the relative rarity of these injuries, surgeons may be less familiar with the distinctive imaging patterns present on advanced imaging as evidenced by lower intraobserver reliability. While the standard procedure for the initial evaluation of a patient presenting with trauma includes anterior-posterior (AP) imaging of the pelvis, only 30% of sacral fractures are appropriately visualized on an AP pelvis radiograph.²² In the absence of a neurological deficit, advanced imaging in the form of computed tomography (CT) is the imaging study of choice. Additionally, local education may also contribute to the differences noted, as surgeons from the regions of Africa, Latin, and South America represent a minority of AO membership demographics. As a result, these regions may be less familiar with the application of AO classification systems in general. Nuanced findings, such as involvement of the S1 facet, may require more experience in the practical application of the classification scheme. Despite aforementioned differences, the overall performance demonstrates substantial reliability across all regions which is comparable to the reliability results of other widely adopted classification systems for spinal trauma.^{11,12,23,24}

To evaluate the need for possible classification refinement based on regional interpretation, the reliability of each individual fracture morphology and subtype was assessed. Those with poor accuracy and reliability may suggest the need for further classification refinement as an inadequately validated system will ultimately serve as a biased predictor of outcomes.²⁵ However, the results of this global validation were promising overall. While no subtypes were found to have low reliability of classification, fracture subtypes C0/1/3 may require particular emphasis when educating observers on the use of the classification system. C1/3 fractures had the lowest agreement percentages for correct classification in 1 or both assessments. Given that there were no significant differences in accuracy between regions for these specific fracture subtypes, the difficulty in classifying such fractures may be inherent to the fracture or based on surgeon experience level as opposed to differences based on regional variability alone.²⁶ Both C0/1 are non-displaced fractures which may be challenging to identify even on advanced imaging. When comparing the first and second assessments for Type C fractures, there was a slight “learning effect” in which interobserver reliability improved during the second assessment in the majority of fracture subtypes, suggesting that deficiencies in performance are able to be overcome with education and repeated practical application.

Although the primary purpose of the study was identification of potential regional variations in the reproducibility and reliability of the AO Spine Sacral Classification System, a previous validation of the AO Spine Upper Cervical Classification System has suggested surgeon experience may affect classification reproducibility.²⁶ Our study identified surgeons with 5–10 and 11–20 years’ experience as the most likely respondents, while surgeons with either less than 5 or greater

than 20 years' experience were less likely to be included. This may partly explain variations in fracture classification reliability and reproducibility as more experienced surgeons may identify nuances in fracture characteristics ultimately affecting fracture subtype classification. Additionally, we identified respondents from North America as significantly more likely to have <5 years of experience, respondents from the Middle East were more likely to have 11–20 or >20 years' experience, while respondents from Africa, Asia, Europe, and Latin/South America had a more evenly distributed number of years in practice. Although outside the scope of this manuscript, additional resources should be aimed at identifying the effect of surgeon experience on fracture classifications.

This study is not without limitations. The surgeons surveyed in this study are all members of AO Spine and/or AO Trauma, and therefore may not represent the true cross section of surgeons globally. While this does impart a selection bias, it afforded the ability to survey the largest group of surgeons managing sacral fractures worldwide. It may be surmised that because of a participation bias toward academic surgeons, the surveyed group may be more familiar with AO classification systems, which can overstate the reliability of the classification system. Lastly, the limited number of participants from some global regions required grouping of regions for statistical comparison reducing the granularity of the data presented. This manifested itself with some regions not having experienced surgeons respond (e.g., North America had no surgeon with >20 years' experience) and other regions still having low overall participation (Africa had 9 total respondents with even distribution of years in practice). While this is suboptimal for a survey study, regional percentages of AO Spine membership are proportionate to our study's participation results underscoring low non-response bias.

Conclusion

In a global validation study of the AO Spine Sacral Classification System, substantial reliability of the classification for both fracture morphology and fracture subtype was found across all geographic regions. The findings of this study underscore the universal nature of the classification scheme without the need for further refinement based upon regional interpretation. These results are essential in laying the foundation for the development of globally accepted classification systems.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.





Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Author's Note

The manuscript submitted does not contain information about medical device(s)/drug(s). This study was organized and funded by AO Spine through the AO Spine Knowledge Forum Trauma, a focused group of international Trauma experts. Relevant financial activities outside the submitted work: consultancy, grants, royalties, stocks, payment for lecture. AO Spine Sacral Classification Group members: Aaron Hilis, Alfredo Guiroy, Aaron Hockley, Ahmed ShawkyAbdelgawaad, Akbar Jaleel Zubairi, Alarcon Olger, Alberto Campione, Alecio Barcelos, Alex del Arco, Alexander Weening, Ali Haghnegahdar, Ali Öner, Alqatub Ahmed, Alsammak Wael, Alvin Pun, Amit Bhandutia, Andjel German, Antonio Sanchez Rodriguez Ashraf El Naga, Barbeiro Gonçalves, Bas Teresa, Bernhard Ullrich, Bernucci Claudio, Bruno Saciloto, Carlos Pernal, Charalampidis Anastasios Christina Cheng, Ciccioli Nicolas, Cumhuri Kilicer, Daniel Cruz, Daniela Linhares, Dare Ashaolu, Das Gurudip, Dave Anthony Dizon, De Falco Giovanni, Dewan Shamsul Asif, Dihan Aponso, Dimitros Patronis, Luis Miguel, Duchén Rodríguez, Elias Enmanuel, Javier Martinez, Enrico Pola, Erol Gercek, Fabii Pozzi, Frank Lyons, Geoffrey Tipper, Glukhov Dmitrii, Gupta Anuj, Hai Nie, Hamouda Waeel, Hany Soliman, Heiller Torres Valencia Huiren Tao, Ichiro Okano, Ignacio Fernández Bances Igor Paredes, Jannardhana Aithala Parampalli, Jason Pittman, Javier Lecaros, Javier Matta, Jeronimo Milano, Joana Freitas, Joana Ganau, Joep Kitzen, Jorge Alves, Jose Joefrey Arbatin, Juan Carlos Ramos Torres Juan Esteban Muñoz Montoya, Juan Lourido, Jun Zhang, Khatri Manoj, Khattab Mohamed Kimmatkar Nitin, Konstantinos Margetis Kumar Vineet, Lady Lozano C, Leandro Badalassi, Lykourgos Kollintzas Maarten de Boer, Manabu Ito, Martin Estefan, Martin Holas, Martinelli Federico, Matias Pereira Duarte, Matti Scholz, Maximo-Alberto Diez-Ulloa Medina Matias, Michel Triffaux, Mohammad El-Sharkawi Morakis Andreas, Mostafa Abdelhafez, Muthu Sathish, Nasser Khan, Nicassio Nicola, Nishanth Ampar, Nuno Neves, Olga Morillo, Osmar Moraes, Osundina Ahmad, Paul van Urk, Pedro Bazán, Pershin Andrey, Peter Ferlic, Peter Loughenbury, , Philippe Bancel, Popescu Eugen Cezar Pragnesh Bhatt, Prakash Sitoula, Przemyslaw Kolakowski Purnajyoti Banerjee, Rabindra Pradhan, Rajesh Bahadur Lakhey Rakesh Kumar, Ramieri Alessandro, Reddy Karthik, Rian Vieira, Ricardo Frada, Ricardo Rodrigues-Pinto Ricky Rasschaert, Russo Fabrizio, Salah Alakkad, Samer Samy, Sbaffi Pier Filippo, Sebastián Anibal Kornfeld Sergio Zylbersztejn, Sicoli Alfredo, Siegmund Lang, Stacey Darwish, Stipe Corluka, Subbiah Jayakumar Subramaniam Macherla haribabu, Sudhir Ganesan, Taha Karimjee, Takuya Nagai, Tarek Elhewala, Tejada Martin, Thami Benzakour, Toluse Adetunji, Tsuyoshi Okudaira, Tulio Rangel, Umit Guler, Vijay Kumar Loya Vishwajeet Singh, Yetkin Söyüncü, Zacharia Silk, Zagorac Slavisa, Zdenek Klezl.

ORCID iDs

Brian A. Karamian, MD  <https://orcid.org/0000-0003-0512-6019>
 Mark J. Lambrechts, MD  <https://orcid.org/0000-0002-9106-2228>
 Jose A. Canseco MD  <https://orcid.org/0000-0002-2152-5725>
 Emiliano N. Vialle  <https://orcid.org/0000-0003-1157-4889>
 Shanmuganathan Rajasekaran  <https://orcid.org/0000-0001-6043-006X>

References

1. Bellabarba C, Stewart JD, Ricci WM, DiPasquale TG, Bohlhofner BR. Midline sagittal sacral fractures in anterior-posterior compression pelvic ring injuries. *J Orthop Trauma*. 2003;17:32-37.
2. Bydon M, De la Garza-Ramos R, Macki M, Desai A, Gokaslan AK, Bydon A. Incidence of sacral fractures and in-hospital postoperative complications in the United States. *Spine*. 2014;39:E1103-E1109.
3. Strange-Vognsen HH, Lebech A. An unusual type of fracture in the upper sacrum. *J Orthop Trauma*. 1991;5:200-203.
4. Denis F, Davis S, Comfort T. Sacral fractures. *Clin Orthop Relat Res*. 1988;227:67-81.
5. Isler B. Lumbosacral lesions associated with pelvic ring injuries. *J Orthop Trauma*. 1990;4:1-6.
6. Roy-Camille R, Saillant G, Gagna G, Mazel C. Transverse fracture of the upper sacrum. *Spine*. 1985;10:838-845.
7. Lehman RA, Kang DG, Bellabarba C. A new classification for complex lumbosacral injuries. *Spine J*. 2012;12:612-628.
8. Bonnin JG. Sacral fractures and cauda equina lesions. *Med World*. 1947;67:140-144.
9. Tile M. Pelvic ring fractures: should they be fixed?. *J Bone Joint Surg Br*. 1988;70-B:1-12.
10. Vaccaro AR, et al. Description and reliability of the AOSpine sacral classification system. *J Bone Joint Surg*. 2020;102(16):1454-1463.
11. Vaccaro AR, Koerner JD, Radcliff KE, et al. AOSpine subaxial cervical spine injury classification system. *Eur Spine J*. 2016;25:2173-2184.
12. Kepler CK, Vaccaro AR, Koerner JD, et al. Reliability analysis of the AOSpine thoracolumbar spine injury classification system by a worldwide group of naïve spinal surgeons. *Eur Spine J*. 2016;25:1082-1086.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
14. Canseco JA, Schroeder GD, Patel PD, et al. Regional and experiential differences in surgeon preference for the treatment of cervical facet injuries: a case study survey with the AO spine cervical classification validation group. *Eur Spine J*. 2020;30:1-7. doi:10.1007/s00586-020-06535-z.
15. Schroeder GD, Kepler CK, Koerner JD, et al. A worldwide analysis of the reliability and perceived importance of an injury to the posterior ligamentous complex in AO type a fractures. *Global Spine J*. 2015;5:378-382.
16. Chhabra HS, Arora M. Neglected traumatic spinal cord injuries: causes, consequences and outcomes in an Indian setting. *Spinal Cord*. 2013;51:238-244.
17. Leiding A, Kim EE, Navarro-Ramirez R, et al. Spinal trauma in Tanzania: current management and outcomes. *J Neurosurg Spine*. 2019;31:103-111.
18. Rathore FA. Neglected traumatic spinal cord injuries: experience sharing from Pakistan. *Spinal Cord*. 2013;51:652-653.
19. Burns AS, O'Connell C. The challenge of spinal cord injury care in the developing world. *J Spinal Cord Med*. 2012;35:3-8.
20. Ngoya PS, Muhogora WE, Pitcher RD. Defining the diagnostic divide: an analysis of registered radiological equipment resources in a low-income African country. *Pan Afr Med*. 2016;25:99.
21. Volpi G. Radiography of diagnostic imaging in Latin America. *Nuclear Medicine and Biomedical Imaging*. 2016;1:10-12.
22. Yi C, Hak DJ. Traumatic spinopelvic dissociation or U-shaped sacral fracture: a review of the literature. *Injury*. 2012;43:402-408.
23. Vaccaro AR, Lehman RA, Hurlbert RJ, et al. A new classification of thoracolumbar injuries. *Spine*. 2005;30:2325-2333.
24. Whang PG, Vaccaro AR, Poelstra KA, et al. The influence of fracture mechanism and morphology on the reliability and validity of two novel thoracolumbar injury classification systems. *Spine*. 2007;32:791-795.
25. Audigé L, Bhandari M, Hanson B, Kellam J. A concept for the validation of fracture classifications. *J Orthop Trauma*. 2005;19:401-406.
26. 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-E270.