

# World Journal of *Orthopedics*

*World J Orthop* 2019 January 18; 10(1): 1-53





### EDITORIAL

- 1 Damage control orthopaedics: State of the art  
*Guerado E, Bertrand ML, Cano JR, Cerván AM, Galán A*

### ORIGINAL ARTICLE

#### Observational Study

- 14 Inter- and intra-rater reliability of vertebral fracture classifications in the Swedish fracture register  
*Morgonsköld D, Warkander V, Savvides P, Wihlborg A, Bouzereau M, Möller H, Gerdhem P*
- 23 Validation of the Danish version of the musculoskeletal tumour society score questionnaire  
*Sæbye CKP, Keller J, Baad-Hansen T*

### CASE REPORT

- 33 Contralateral trapezius transfer to treat scapular winging: A case report and review of literature  
*Gil-Álvarez JJ, García-Parra P, Anaya-Rojas M, Martínez-Fuentes MDP*
- 45 Peroneus longus tendon rupture: A case report  
*Koh D, Liow L, Cheah J, Koo K*

**ABOUT COVER**

Editor-in-chief of *World Journal of Orthopedics*, Baogan Peng, MD, PhD, Chief Doctor, Professor, Department of Spinal Surgery, General Hospital of Armed Police Force, Beijing 100039, China

**AIMS AND SCOPE**

*World Journal of Orthopedics* (*World J Orthop*, *WJO*, online ISSN 2218-5836, DOI: 10.5312) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians.

*WJO* covers topics concerning arthroscopy, evidence-based medicine, epidemiology, nursing, sports medicine, therapy of bone and spinal diseases, bone trauma, osteoarthritis, bone tumors and osteoporosis, etc. Priority publication will be given to articles concerning diagnosis and treatment of orthopedic diseases. The following aspects are covered: Clinical diagnosis, laboratory diagnosis, differential diagnosis, imaging tests, pathological diagnosis, molecular biological diagnosis, immunological diagnosis, genetic diagnosis, etc.

We encourage authors to submit their manuscripts to *WJO*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

**INDEXING/ABSTRACTING**

*World Journal of Orthopedics* (*WJO*) is now abstracted and indexed in PubMed, PubMed Central, Emerging Sources Citation Index (Web of Science), China National Knowledge Infrastructure (CNKI), China Science and Technology Journal Database (CSTJ), and Superstar Journals Database.

**RESPONSIBLE EDITORS  
FOR THIS ISSUE**

Responsible Electronic Editor: *Wen-Wen Tan*      Proofing Editorial Office Director: *Jin-Lei Wang*

**NAME OF JOURNAL**

*World Journal of Orthopedics*

**ISSN**

ISSN 2218-5836 (online)

**LAUNCH DATE**

November 18, 2010

**FREQUENCY**

Irregular

**EDITORS-IN-CHIEF**

Baogan Peng

**EDITORIAL BOARD MEMBERS**

<http://www.wjgnet.com/2218-5836/editorialboard.htm>

**EDITORIAL OFFICE**

Jin-Lei Wang, Director

**PUBLICATION DATE**

January 18, 2019

**COPYRIGHT**

© 2019 Baishideng Publishing Group Inc

**INSTRUCTIONS TO AUTHORS**

<https://www.wjgnet.com/bpg/gerinfo/204>

**GUIDELINES FOR ETHICS DOCUMENTS**

<https://www.wjgnet.com/bpg/GerInfo/287>

**GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH**

<https://www.wjgnet.com/bpg/gerinfo/240>

**PUBLICATION MISCONDUCT**

<https://www.wjgnet.com/bpg/gerinfo/208>

**ARTICLE PROCESSING CHARGE**

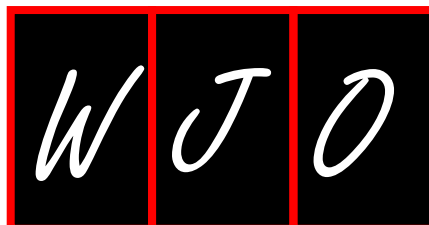
<https://www.wjgnet.com/bpg/gerinfo/242>

**STEPS FOR SUBMITTING MANUSCRIPTS**

<https://www.wjgnet.com/bpg/GerInfo/239>

**ONLINE SUBMISSION**

<https://www.f6publishing.com>



## Observational Study

# Inter- and intra-rater reliability of vertebral fracture classifications in the Swedish fracture register

David Morgonsköld, Victoria Warkander, Panayiotis Savvides, Axel Wihlborg, Mathilde Bouzereau, Hans Möller, Paul Gerdhem

**ORCID number:** David Morgonsköld (0000-0002-7248-9488); Victoria Warkander (0000-0002-7516-8831); Panayiotis Savvides (0000-0003-3580-0334); Axel Wihlborg (0000-0002-5355-9655); Mathilde Bouzereau (0000-0003-4570-1815); Hans Möller (0000-0003-2339-6274); Paul Gerdhem (0000-0001-8061-7163).

**Author contributions:** Morgonsköld D, Warkander V, Möller H and Gerdhem P designed the study. All of the authors collected the data. Morgonsköld D, Warkander V and Gerdhem P made the analysis and manuscript draft. Savvides P, Wihlborg A, Bouzereau M, Möller H and Gerdhem P made the manuscript comment. Morgonsköld D and Gerdhem P finalized the manuscript.

### Institutional review board

**statement:** The regional Ethical Review Board in Stockholm, Sweden, approved the study, No. 2016/897-31/1. All patients were investigated and treated according to clinical guidelines. No interventions were made.

### Informed consent statement:

Patient consent was not required for the retrieval of radiological images from the hospital. Verbal consent was obtained from the physicians that classified the images.

### Conflict-of-interest statement:

None of the authors has any conflicting interests.

**David Morgonsköld, Panayiotis Savvides, Axel Wihlborg, Hans Möller, Paul Gerdhem,** Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm SE-14186, Sweden

**Victoria Warkander,** Department of Orthopaedics, Danderyd Hospital, Danderyd 18288, Sweden

**Panayiotis Savvides, Axel Wihlborg, Hans Möller, Paul Gerdhem,** Department of Orthopaedics, Karolinska University Hospital, Stockholm SE-14186, Sweden

**Mathilde Bouzereau,** Department of Emergency Medicine, Karolinska University Hospital, Stockholm SE-14186, Sweden

**Corresponding author:** Paul Gerdhem, MD, PhD, Associate Professor, Department of Orthopaedics, Karolinska University Hospital, K54, Huddinge, Stockholm SE-14186, Sweden. [paul.gerdhem@sll.se](mailto:paul.gerdhem@sll.se)  
**Telephone:** +46-58-580000

## Abstract

### AIM

To investigate the inter- and intra-rater reliability of the vertebral fracture classifications used in the Swedish fracture register.

### METHODS

Radiological images of consecutive patients with cervical spine fractures ( $n = 50$ ) were classified by 5 raters with different experience levels at two occasions. An identical process was performed with thoracolumbar fractures ( $n = 50$ ). Cohen's kappa was used to calculate the inter- and intra-rater reliability.

### RESULTS

The mean kappa coefficient for inter-rater reliability ranged between 0.54 and 0.79 for the cervical fracture classifications, between 0.51 and 0.72 for the thoracolumbar classifications (overall and for different sub classifications), and between 0.65 and 0.77 for the presence or absence of signs of ankylosing disorder in the fracture area. The mean kappa coefficient for intra-rater reliability ranged between 0.58 and 0.80 for the cervical fracture classifications, between 0.46 and 0.68 for the thoracolumbar fracture classifications (overall and for different sub classifications) and between 0.79 and 0.81 for the presence or absence of signs of ankylosing disorder in the fracture area.



**Data sharing statement:** Data can be provided on request by the corresponding author.

**Open-Access:** This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

**Manuscript source:** Unsolicited manuscript

**Received:** September 12, 2018

**Peer-review started:** September 12, 2018

**First decision:** October 15, 2018

**Revised:** October 27, 2018

**Accepted:** December 24, 2018

**Article in press:** December 24, 2018

**Published online:** January 18, 2019

## CONCLUSION

The classifications used in the Swedish fracture register for vertebral fractures have an acceptable inter- and intra-rater reliability with a moderate strength of agreement.

**Key words:** Vertebral; Spine; Fracture; Classification; Swedish fracture register; Inter-rater; Intra-rater; Reliability

©The Author(s) 2019. Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** The Swedish Fracture register gathers national data on fractures. We adapted commonly used classifications for spine fractures and studied inter and intra-rater reliability as a basis for future usage of the register, including research on outcome after spine fractures.

**Citation:** Morgonsköld D, Warkander V, Savvides P, Wihlborg A, Bouzereau M, Möller H, Gerdhem P. Inter- and intra-rater reliability of vertebral fracture classifications in the Swedish fracture register. *World J Orthop* 2019; 10(1): 14-22  
**URL:** <https://www.wjnet.com/2218-5836/full/v10/i1/14.htm>  
**DOI:** <https://dx.doi.org/10.5312/wjo.v10.i1.14>

## INTRODUCTION

One in six women and one in twelve men will at some point in their lifetime sustain a vertebral fracture<sup>[1]</sup>. The current knowledge of the epidemiology of vertebral fractures in Sweden is based on the national patient register, which only classifies fractures based on large anatomical areas<sup>[2,3]</sup>. Classification of vertebral fractures is important to describe epidemiology and allow treatment comparisons. By using large quality registers, we have the possibility to improve knowledge, increase our understanding of vertebral fracture epidemiology and possibly have an impact on future fracture treatment. The recently introduced Swedish Fracture Register offers such an opportunity<sup>[4]</sup>.

Most of the previous studies on inter-rater and/or intra-rater classification reliability have been performed in expert settings<sup>[5-10]</sup>, while fewer studies have been performed in more clinical settings with physicians of different experience levels<sup>[11,12]</sup>. Registration of vertebral fractures was recently introduced in the Swedish Fracture Register. The reliability of the classification system has not been tested. Our aim was therefore to test the classification before large scale epidemiologic or observational studies in this register are attempted. Our hypotheses were that inter- and intra-rater reliability in this register are acceptable and similar to previously published studies.

## MATERIALS AND METHODS

### Fracture classification

Vertebral fractures have been possible to enter into the Swedish Fracture Register since 2015<sup>[13]</sup>. The fracture classification is performed with the help of drawings that provide information about typical fracture patterns. Since classifications were to be performed not only by experts in spinal trauma care, but also by interns, residents and emergency physicians, the aim was to use classifications that was easy to comprehend without previous experience in the subject. Earlier published classifications were used as a foundation, but some of them were modified compared to their original descriptions. Appendix I (available as supplementary material) show the fracture classifications used in the Register.

Atlas (C1) fractures are classified according to Jackson<sup>[14]</sup>. Axis (C2) fractures are classified according to Anderson-D'Alonzo<sup>[15]</sup> for odontoid fractures and according to Effendi<sup>[16]</sup> and Levine-Edwards<sup>[17]</sup> for traumatic spondylolisthesis of the axis (hangman's fracture). Subaxial (C3-Th1) fractures are classified with a slightly modified version of the sub-axial cervical spine injury classification (SLIC)<sup>[18]</sup>. Thoracic and lumbar (Th2-L5) fractures are classified with a modified version of the 2013 version of the "Arbeitsgemeinschaft für Osteosynthesefragen" (AO) spine injury

classification system<sup>[19]</sup>.

In addition, the classifying physician is also asked to determine whether there are signs of ankylosing spinal disorders in the fracture area, such as diffuse idiopathic skeletal hyperostosis (DISH) or ankylosing spondylitis (AS).

### **Study population**

We identified consecutive patients with vertebral fractures, 50 with cervical spine fractures and 50 with thoracolumbar spine fractures, from the medical records at the Karolinska University Hospital Huddinge. It is the primary hospital for the southern part of Stockholm, but also contains one of two referral centers for surgical treatment of vertebral fractures in the Stockholm County, which has a total population of 2.3 million inhabitants. The radiological images (plane x-ray, computed tomography, and/or magnetic resonance imaging; MRI) were collected from the radiological archive. The mean age at the time of the radiological examination was 60 years (range 15-93) for patients with cervical fractures and 55 years (range 9-96) for patients with thoracolumbar fractures. Low-energy trauma (defined as a fall from standing height or less) was the cause for 23 (46%) of the cervical fractures and 20 (40%) of the thoracolumbar fractures. Non-surgical treatment was provided to 40 (80%) of the patients with cervical fractures and to 35 (70%) of the patients with thoracolumbar fractures.

### **Reliability tests**

In all, six physicians of different experience level and one medical student were involved in the classifications; five of which were involved in each of the cervical and the thoracolumbar classifications with two of the physicians and the medical student involved in both classifications (Table 1). The raters only had information on patient's age and date of the investigation. No other clinical information or information about the fractures, such as radiological image evaluations, was provided to the raters. The raters classified the fractures independently without knowledge of the results of the others. The classifications were made with a document containing the Swedish Fracture Register fracture classification and subheadings for each fracture type (Appendix I).

### **Gold standard**

After the classifications were completed, the results from the first test of the two most experienced raters, (Rater 1 and 2) were compared. In case a classification was not identical, Rater 1 and 2 reviewed the available radiological images (plain radiographs, computed tomography and/or magnetic resonance images) until consensus was achieved.

### **Statistics**

The classifications made by the raters were analyzed using Cohen's kappa<sup>[20]</sup> and are shown as kappa value (asymptotic standard error, standard deviation or range). Besides analyzing the result of all cervical and all thoracolumbar classifications, the classifications made by the raters were divided into subgroups analyzing only C1, C2 or C3-Th1 injuries in the cervical spine and A-, B- or C-types of injuries in the thoracolumbar spine (Th2-L5). All statistical analyzes were made in IBM SPSS Statistics for Windows, Version 24.0.

## **RESULTS**

### **Gold standard**

Rater 1 and 2 reached independent consensus in 21 patients with cervical fractures and 20 patients with thoracolumbar fractures. For presence or no presence of ankylosing spinal disorders, independent consensus was reached in 45 patients with cervical fractures and 46 patients with thoracolumbar fractures. After discussion, final consensus was reached for all fracture classifications and signs of ankylosing spinal disorder. The final distribution of fracture classifications is presented in Table 2. Not all classifications were represented.

### **Inter-rater reliability for vertebral fracture classifications**

Inter-rater reliability mean kappa was 0.54 (0.26-0.70) for the cervical fracture classifications (C1-Th1) and 0.51 (0.18-0.69) for the thoracolumbar fracture classification (Th2-L5). Inter-rater reliability for individual raters is presented in Table 3.

For the cervical subgroups the inter-rater reliability was 0.79 (0.61-1.00) for C1 fractures, 0.76 (0.56-0.89) for C2 fractures and 0.68 (0.45-0.80) for C3-Th1 fractures.

**Table 1** Rater experience level and time between the test occasions

	Experience level	Time between first and second test for cervical fractures	Time between first and second test for thoracolumbar fractures
Rater 1	Senior consultant in orthopaedics and spine surgery as well as responsible for the vertebral fracture classification in the Swedish Fracture Register.	1 mo	7 mo
Rater 2	Senior consultant in orthopaedics and spine surgery.	1 mo	7 mo
Rater 3	Specialist in orthopaedics.	7 mo	
Rater 4	Resident in orthopaedics.	7 mo	
Rater 5	Medical student, trained by Rater 1 to be knowledgeable on the topic.	1 mo	1 mo
Rater 6	Resident in orthopaedics.		1 yr
Rater 7	Resident in emergency medicine.		1 yr

Raters 1, 2 and 5 were actively working with classifications and vertebral fractures in the time between the tests while Raters 3, 4, 6 and 7 were not.

Inter-rater reliability for the thoracolumbar classifications was 0.72 (0.34-0.88) for A-type of injuries, 0.60 (0.41-0.76) for B-type injuries and 0.70 (0.51-0.89) for C-type of injuries.

#### ***Signs of ankylosing spinal disorder inter-rater reliability***

For presence or absence of signs of ankylosing spinal disorder the inter-rater mean kappa coefficient was 0.65 (0.20-0.91) for cervical fractures and 0.77 (0.31-1.00) for thoracolumbar fractures. Inter-rater reliability for individual raters is presented in [Table 3](#).

#### ***Intra-rater reliability for vertebral fracture classifications***

Intra-rater reliability mean kappa was 0.58 (0.40-0.72) for cervical fracture classifications (C1-Th1) and 0.46 (0.16-0.62) for the thoracolumbar fracture classification (Th2-L5). Intra-rater reliability for individual raters is presented in [Table 4](#).

For the cervical subgroups the intra-rater reliability was 0.80 (0.67-0.93) for C1 fractures, 0.78 (0.63-0.93) for C2 fractures and 0.69 (0.53-0.82) for C3-Th1 fractures. Intra-rater reliability mean kappa for the simplified thoracolumbar classifications was 0.63 (0.31-0.83) for only A-type of injuries, 0.56 (0.32-0.78) for only B-type injuries and 0.68 (0.47-0.90) for only C-type of injuries.

#### ***Signs of ankylosing spinal disorder intra-rater reliability***

For presence or absence of signs of ankylosing spinal disorder the intra-rater mean kappa was 0.79 (0.66-1.00) for the cervical spine and 0.81 (0.60-1.00) for the thoracolumbar spine. Intra-rater reliability for individual raters is presented in [Table 4](#).

## **DISCUSSION**

According to Landis and Koch<sup>[21]</sup>, the overall inter- and intra-rater reliability had a moderate strength of agreement for both cervical and thoracolumbar fracture classifications. When dividing the classifications into subgroups, both inter- and intra-rater reliability mean kappa coefficients increased. Inter- and intra-rater kappa coefficients were also generally higher for the more experienced physicians and the medical student who received training, which was most evident in the thoracolumbar classification resulting in the largest variability in inter-rater and intra-rater Kappa values. Similarly, our overall results for signs of ankylosing spinal disorder inter-rater reliability had a substantial strength of agreement while the intra-rater reliability have a substantial to almost perfect strength of agreement for vertebral fractures<sup>[21]</sup>.

Previous studies on fracture classifications used in the Swedish Fracture Register include humerus and tibia fracture classifications that have produced results similar to or better than our overall inter- and intra-rater reliability for the cervical and thoracolumbar classifications<sup>[5,9]</sup>. Both studies used only one classification system and had a different study design. The humerus classifications had 12 possible alternatives

**Table 2** Distribution of fractures according to the gold standard

			Number of patients	Number of patients with signs of ankylosing spinal disorder
Cervical spine	C1-Th1		50	6
	C1	Posterior arch	2	0
		Burst	5	1
		Lateral mass	1	0
	C2	Odontoid	11	1
		Hangman's	2	0
		Unclassifiable	4	0
	C3-Th1	Compression	13	3
		Burst	0	0
		Translation/rotation	7	2
		Other	11	0
		Unclassifiable	2	0
Thoracolumbar spine	Th2-L5	Injury through axial compression (A-type injuries)	40	6
		Injury through anterior and posterior structures causing displacement (C-type injuries)	10	10
	Th2-L5 A-type injuries	Wedge shaped compressions	17	2
		Fracture through the middle part of the vertebral body	2	0
		Burst fracture	21	4
	Th2-L5 B-type injuries <sup>1</sup>	No injury of posterior structures	24	1
		Rupture of the posterior tension band structures through bone	9	4
		Rupture of the posterior ligament	5	0
		Unable to assess	2	1
	Th2-L5 C-type injuries	Hyperextension injury without translation	9	9
		Translation injury/dislocation through bone or disc/ligament	1	1

<sup>1</sup>Classification of B-type injuries are only assessed in patients with A-type injuries. C: Cervical vertebra; L: Lumbar vertebra; Th: Thoracic vertebra.

while the tibia classifications had 27 possible alternatives for the raters to choose from<sup>[5,9]</sup>. This can be compared to our study using 5 different classification systems with 3-14 possible alternatives in each separate classification system (in total 33 alternatives).

One study on C2 odontoid classifications and another study on the AO Spine subaxial cervical injury system (C3-7) have produced slightly poorer inter-rater reliability when compared to our subgroups (C2 and C3-Th1)<sup>[7,8]</sup>. Previous studies on the AO Spine thoracolumbar spine injury classification system and Thoracolumbar Injury Classification and Severity Score (TLICS) have produced results that are generally better than our results for the Th2-L5 fracture classification<sup>[6,11,19]</sup>. However, two more recently published studies concerning the AO Spine thoracolumbar spine injury classification system had poorer or similar inter-rater reliability compared to our study<sup>[10,12]</sup>. The differences in results between studies are most likely due to differences in methods, study populations and the limitations in using Kappa, but we interpret our inter-rater results for vertebral fracture classifications used in the Swedish Fracture Register to be within what others have previously considered acceptable for a classification system<sup>[5,7-10,12]</sup>.

Our results for intra-rater reliability was at best similar to that of some previous studies<sup>[5-10]</sup>, while similar to or better than the results from a study with raters of

**Table 3 Inter-rater reliability compared to the gold standard shown as kappa value (asymptotic standard error) for vertebral fracture classifications**

		Kappa									
		Cervical spine					Thoracolumbar spine				
		C1-Th1	C1	C2	C3-Th1	Signs of ankylosing spinal disorder <sup>1</sup>	Th2-L5	Th2-L5 A-type injuries	Th2-L5 B-type injuries	Th2-L5 C-type injuries	Signs of ankylosing spinal disorder <sup>1</sup>
Test 1, Rater	1	0.59 (0.07)	0.61 (0.10)	0.81 (0.07)	0.78 (0.06)	0.81 (0.13)	0.64 (0.08)	0.88 (0.06)	0.70 (0.08)	0.72 (0.10)	0.91 (0.07)
	2	0.59 (0.07)	0.72 (0.10)	0.89 (0.06)	0.71 (0.07)	0.91 (0.09)	0.67 (0.07)	0.86 (0.06)	0.64 (0.08)	0.89 (0.08)	1.00 (0.00)
	3	0.51 (0.07)	0.79 (0.10)	0.66 (0.08)	0.66 (0.07)	0.67 (0.15)					
	4	0.47 (0.07)	0.86 (0.09)	0.74 (0.07)	0.61 (0.07)	0.32 (0.11)					
	5	0.59 (0.07)	0.93 (0.07)	0.74 (0.07)	0.73 (0.07)	0.77 (0.13)	0.44 (0.08)	0.67 (0.08)	0.63 (0.09)	0.57 (0.11)	0.91 (0.07)
	6						0.46 (0.08)	0.76 (0.08)	0.49 (0.09)	0.81 (0.09)	0.75 (0.10)
	7						0.45 (0.08)	0.67 (0.08)	0.48 (0.09)	0.61 (0.13)	0.31 (0.13)
Test 1 Mean		0.55 (0.06)	0.78 (0.12)	0.77 (0.08)	0.70 (0.07)	0.70 (0.23)	0.53 (0.11)	0.77 (0.10)	0.59 (0.10)	0.72 (0.13)	0.77 (0.27)
Test 2, Rater	1	0.70 (0.07)	0.93 (0.07)	0.81 (0.07)	0.80 (0.06)	0.81 (0.13)	0.58 (0.08)	0.82 (0.07)	0.76 (0.07)	0.77 (0.09)	0.91 (0.07)
	2	0.65 (0.07)	0.65 (0.10)	0.89 (0.06)	0.78 (0.06)	0.77 (0.13)	0.60 (0.08)	0.82 (0.07)	0.62 (0.08)	0.78 (0.10)	0.86 (0.08)
	3	0.26 (0.06)	0.66 (0.09)	0.56 (0.07)	0.45 (0.07)	0.52 (0.14)					
	4	0.44 (0.07)	0.72 (0.10)	0.65 (0.08)	0.63 (0.07)	0.20 (0.08)					
	5	0.63 (0.07)	1.00 (0.00)	0.85 (0.07)	0.70 (0.07)	0.74 (0.15)	0.69 (0.07)	0.82 (0.07)	0.76 (0.07)	0.82 (0.10)	0.95 (0.05)
	6						0.35 (0.08)	0.58 (0.09)	0.46 (0.09)	0.51 (0.13)	0.70 (0.11)
	7						0.18 (0.07)	0.34 (0.09)	0.41 (0.10)	0.52 (0.14)	0.38 (0.13)
Test 2 Mean		0.54 (0.18)	0.79 (0.16)	0.75 (0.14)	0.67 (0.14)	0.61 (0.25)	0.48 (0.21)	0.62 (0.21)	0.60 (0.17)	0.68 (0.15)	0.76 (0.23)

For test 1 and 2 mean kappa (standard deviation) is shown. <sup>1</sup>Ankylosing spondylitis or diffuse idiopathic skeletal hyperostosis. C: Cervical vertebra; L: Lumbar vertebra; Th: Thoracic vertebra.

different experience levels<sup>[12]</sup>. The extended time (up to a year) in between the two reliability tests in our study needs to be taken into consideration, as well as lack of specific training and lack of detailed data on injury and patient history. Most other studies have about 4-6 wk between tests<sup>[5,7,9,10,12]</sup>, with the longer times being 6-8 mo<sup>[6,8]</sup>. Some studies have also provided more information about the patients to the raters such as; clinical information (patient history, injury mechanism, associated injuries, clinical examination and neurological status), operation notes, follow up information, fracture level or radiological image evaluations<sup>[6,7,9-11,19]</sup>. Therefore, our study could be said to represent the minimum level, which could have affected both our inter- and intra-rater reliability results negatively compared to other studies that provide more information.

In the discussions to decide a gold standard some of the common problems that came up were disagreement whether the example image, description text or the assumed injury mechanism should take priority. Classifications that involve injury to ligaments also created uncertainty among the raters in cases in which no MRI were provided. Still MRI is generally not necessary for thoracolumbar fractures<sup>[22]</sup> and is not always available for cervical fractures. From the results it is apparent that thoracolumbar B-type injuries may be the most difficult to classify (Table 2). This is not surprising since this corresponds to our clinical experience; to determine whether there is a rupture of the posterior ligament complex is not always easy.

It has also previously been suggested that descriptions of the mechanisms of injury and ligamentous injury should not be included in a spinal injury classification<sup>[23]</sup>. Yet it must also be taken into the consideration that the classifications need to be clinically relevant and must be associated with relevant patient outcomes in context with specific fracture management plans<sup>[23,24]</sup>.

The subgroups and simplified classification with fewer alternatives have a higher mean kappa coefficient for both inter- and intra-rater reliability, the most obvious reason for this increase is the limitations in using Cohen's kappa as it does not take in consideration if the difference between two classifications is small or large. Another possible reason for the increase is that more choices make it harder to choose the right classification, especially if they are similar to each other. Improved descriptions could



**Table 4** Intra-rater reliability for vertebral fracture classifications shown as kappa value (asymptotic standard error) for all raters. Raters are compared to themselves.

		Kappa									
		Cervical spine					Thoracolumbar spine				
		C1-Th1	C1	C2	C3-Th1	Signs of ankylosing spinal disorder <sup>1</sup>	Th2-L5	Th2-L5 A-type injuries	Th2-L5 B-type injuries	Th2-L5 C-type injuries	Signs of ankylosing spinal disorder <sup>1</sup>
Rater	1	0.72 (0.07)	0.67 (0.10)	0.93 (0.05)	0.80 (0.06)	1.00 (0.00)	0.59 (0.08)	0.76 (0.08)	0.69 (0.08)	0.84 (0.09)	1.00 (0.00)
	2	0.62 (0.07)	0.72 (0.10)	0.93 (0.05)	0.75 (0.07)	0.70 (0.14)	0.62 (0.07)	0.83 (0.07)	0.64 (0.08)	0.90 (0.07)	0.86 (0.08)
	3	0.40 (0.07)	0.86 (0.09)	0.63 (0.07)	0.55 (0.07)	0.66 (0.12)					
	4	0.44 (0.07)	0.79 (0.10)	0.64 (0.08)	0.53 (0.07)	0.72 (0.10)					
	5	0.72 (0.07)	0.93 (0.07)	0.78 (0.07)	0.82 (0.06)	0.85 (0.10)	0.60 (0.08)	0.75 (0.07)	0.78 (0.08)	0.69 (0.10)	0.95 (0.05)
	6						0.32 (0.07)	0.52 (0.09)	0.37 (0.10)	0.49 (0.14)	0.60 (0.13)
	7						0.16 (0.07)	0.31 (0.09)	0.32 (0.10)	0.47 (0.17)	0.63 (0.19)

<sup>1</sup>Ankylosing spondylitis or diffuse idiopathic skeletal hyperostosis. C: Cervical vertebra; L: Lumbar vertebra; Th: Thoracic vertebra.

possibly improve the classifications. Reduced number of classifications could result in greater agreement between observers, but will decrease details of fracture data. Implementing web based training/example videos could possibly increase the agreement between physicians.

### Strength and limitations

There are currently no methodological standard<sup>[25]</sup> and we chose to use Cohen's kappa in our study as it is commonly used in other similar studies<sup>[5,8,9,12,25]</sup>. A clinical setting with raters of different experience levels classifying vertebral fractures is similar to how data collection of vertebral fractures to the Swedish Fracture Register works and thus, this study provides us with an insight on the reliability of the data in the Swedish Fracture Register. Unfortunately this also make comparisons to other studies harder as different methods, study populations, classification systems and/or comparisons for the classification system is used<sup>[6,7,11,19]</sup>. From this, we can assess that our results are not necessarily applicable to other registers or studies using one or more of the five individual classification systems used in the Swedish Fracture Register. The study population in this study cannot be generalized for the entire population or every type of hospital. It can also be argued that the study population in this study is too small for the cervical spine and that 50 patients should be selected for each classification system within the cervical spine in order to get a better understanding of the inter- and intra-rater reliability for these individual classification systems. Nevertheless, it would be hard to find large numbers of cervical fractures for a similar study, especially C1 fractures, which only represents 2% of all vertebral fractures<sup>[4]</sup> but could be the focus for future studies.

### Conclusions

The classifications used in the Swedish Fracture Register for vertebral fractures have an acceptable inter- and intra-rater reliability with a moderate strength of agreement. With specific training or many years of experience, a higher consistency in inter-rater reliability could be achieved. The results indicate that the Swedish Fracture Register data may be used for studies on epidemiology. Studies comparing treatment outcome should consider reclassifying images to ensure correct classification.

## ARTICLE HIGHLIGHTS

### Research background

The Swedish Fracture register gives the possibility to attain nationwide data on fractures.

### Research motivation

Classification of vertebral fractures have recently been introduced in the register.

### Research objectives

We tested the inter- and intra-rater reliability of the vertebral fracture classifications in the

Swedish Fracture register.

### Research methods

Radiological images of consecutive patients with vertebral fractures were classified by 5 raters with different experience levels at two occasions.

### Research results

The mean kappa coefficient for inter-rater reliability ranged between 0.51 and 0.79 for the different classifications (overall and for different sub classifications). The mean kappa coefficient for intra-rater reliability ranged between 0.46 and 0.81.

### Research conclusions

The classifications have an acceptable inter- and intra-rater reliability with a moderate strength of agreement.

### Research perspectives

The results indicate that the Swedish Fracture register is ready for use in epidemiological studies. Studies comparing treatment outcome should consider reclassifying images to ensure correct classification.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the collaboration of Michael Möller and Carl Ekholm, initiators of the Swedish Fracture Register, and Helena Brisby, Sahlgrenska University Hospital and Björn Knutsson, Sundsvall Härnösand Hospital for the discussions and implementations of the vertebral fracture classifications in the Register.

## REFERENCES

- 1 Gerdhem P. Osteoporosis and fragility fractures: Vertebral fractures. *Best Pract Res Clin Rheumatol* 2013; **27**: 743-755 [PMID: 24836333 DOI: 10.1016/j.berh.2014.01.002]
- 2 Jansson KA, Blomqvist P, Svedmark P, Granath F, Buskens E, Larsson M, Adami J. Thoracolumbar vertebral fractures in Sweden: an analysis of 13,496 patients admitted to hospital. *Eur J Epidemiol* 2010; **25**: 431-437 [PMID: 20449637 DOI: 10.1007/s10654-010-9461-5]
- 3 The National Board of Health and Welfare in Sweden National Patient Registry. [Accessed 10 Dec 2016]. Available from: <http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/inenglish>
- 4 Svenska Frakturregistret Årsrapport 2016. Available from: <http://www.frakturregistret.se>
- 5 Wennergren D, Stjernström S, Möller M, Sundfeldt M, Ekholm C. Validity of humerus fracture classification in the Swedish fracture register. *BMC Musculoskelet Disord* 2017; **18**: 251 [PMID: 28601085 DOI: 10.1186/s12891-017-1612-3]
- 6 Lewkonja P, Paolucci EO, Thomas K. Reliability of the thoracolumbar injury classification and severity score and comparison with the denis classification for injury to the thoracic and lumbar spine. *Spine (Phila Pa 1976)* 2012; **37**: 2161-2167 [PMID: 22648029 DOI: 10.1097/BRS.0b013e3182601469]
- 7 Silva OT, Sabba MF, Lira HI, Ghizoni E, Tedeschi H, Patel AA, Joaquim AF. Evaluation of the reliability and validity of the newer AOSpine subaxial cervical injury classification (C-3 to C-7). *J Neurosurg Spine* 2016; **25**: 303-308 [PMID: 27104288 DOI: 10.3171/2016.2.SPINE151039]
- 8 Barker L, Anderson J, Chesnut R, Nesbit G, Tjauw T, Hart R. Reliability and reproducibility of dens fracture classification with use of plain radiography and reformatted computer-aided tomography. *J Bone Joint Surg Am* 2006; **88**: 106-112 [PMID: 16391255 DOI: 10.2106/JBJS.D.02834]
- 9 Wennergren D, Ekholm C, Sundfeldt M, Karlsson J, Bhandari M, Möller M. High reliability in classification of tibia fractures in the Swedish Fracture Register. *Injury* 2016; **47**: 478-482 [PMID: 26620119 DOI: 10.1016/j.injury.2015.11.002]
- 10 Kaul R, Chhabra HS, Vaccaro AR, Abel R, Tuli S, Shetty AP, Das KD, Mohapatra B, Nanda A, Sangondimath GM, Bansal ML, Patel N. Reliability assessment of AOSpine thoracolumbar spine injury classification system and Thoracolumbar Injury Classification and Severity Score (TLICS) for thoracolumbar spine injuries: results of a multicentre study. *Eur Spine J* 2017; **26**: 1470-1476 [PMID: 27334493 DOI: 10.1007/s00586-016-4663-5]
- 11 Savage JW, Moore TA, Arnold PM, Thakur N, Hsu WK, Patel AA, McCarthy K, Schroeder GD, Vaccaro AR, Dimar JR, Anderson PA. The Reliability and Validity of the Thoracolumbar Injury Classification System in Pediatric Spine Trauma. *Spine (Phila Pa 1976)* 2015; **40**: E1014-E1018 [PMID: 26426715 DOI: 10.1097/BRS.0000000000001011]
- 12 Cheng J, Liu P, Sun D, Qin T, Ma Z, Liu J. Reliability and reproducibility analysis of the AOSpine thoracolumbar spine injury classification system by Chinese spinal surgeons. *Eur Spine J* 2017; **26**: 1477-1482 [PMID: 27807778 DOI: 10.1007/s00586-016-4842-4]
- 13 Wennergren D, Ekholm C, Sandelin A, Möller M. The Swedish fracture register: 103,000 fractures registered. *BMC Musculoskelet Disord* 2015; **16**: 338 [PMID: 26546157 DOI: 10.1186/s12891-015-0795-8]
- 14 Jackson RS, Banit DM, Rhyne AL 3rd, Darden BV 2nd. Upper cervical spine injuries. *J Am Acad Orthop Surg* 2002; **10**: 271-280 [PMID: 15089076 DOI: 10.5435/00124635-200207000-00005]
- 15 Anderson LD, D'Alonzo RT. Fractures of the odontoid process of the axis. 1974. *J Bone Joint Surg Am* 2004; **86-A**: 2081 [PMID: 15342776 DOI: 10.2106/00004623-197456080-00017]

- 16 **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**: 319-327 [PMID: 7263741 DOI: [10.1302/0301-620X.63B3.7263741](https://doi.org/10.1302/0301-620X.63B3.7263741)]
- 17 **Levine AM**, Edwards CC. The management of traumatic spondylolisthesis of the axis. *J Bone Joint Surg Am* 1985; **67**: 217-226 [PMID: 3968113 DOI: [10.2106/00004623-198567020-00007](https://doi.org/10.2106/00004623-198567020-00007)]
- 18 **Vaccaro AR**, Hulbert RJ, Patel AA, Fisher C, Dvorak M, Lehman RA Jr, Anderson P, Harrop J, Oner FC, Arnold P, Fehlings M, Hedlund R, Madrazo I, Rechtine G, Aarabi B, Shainline M; Spine Trauma Study Group. The subaxial cervical spine injury classification system: a novel approach to recognize the importance of morphology, neurology, and integrity of the disco-ligamentous complex. *Spine (Phila Pa 1976)* 2007; **32**: 2365-2374 [PMID: 17906580 DOI: [10.1097/BRS.0b013e3181557b92](https://doi.org/10.1097/BRS.0b013e3181557b92)]
- 19 **Reinhold M**, Audigé L, Schnake KJ, Bellabarba C, Dai LY, Oner FC. AO spine injury classification system: a revision proposal for the thoracic and lumbar spine. *Eur Spine J* 2013; **22**: 2184-2201 [PMID: 23508335 DOI: [10.1007/s00586-013-2738-0](https://doi.org/10.1007/s00586-013-2738-0)]
- 20 **Cohen J**. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; **20**: 37-46 [DOI: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)]
- 21 **Landis JR**, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159-174 [PMID: 843571 DOI: [10.2307/2529310](https://doi.org/10.2307/2529310)]
- 22 **Rajasekaran S**, Vaccaro AR, Kanna RM, Schroeder GD, Oner FC, Vialle L, Chapman J, Dvorak M, Fehlings M, Shetty AP, Schnake K, Maheshwaran A, Kandziora F. The value of CT and MRI in the classification and surgical decision-making among spine surgeons in thoracolumbar spinal injuries. *Eur Spine J* 2017; **26**: 1463-1469 [PMID: 27250728 DOI: [10.1007/s00586-016-4623-0](https://doi.org/10.1007/s00586-016-4623-0)]
- 23 **van Middendorp JJ**, Audigé L, Hanson B, Chapman JR, Hosman AJ. What should an ideal spinal injury classification system consist of? A methodological review and conceptual proposal for future classifications. *Eur Spine J* 2010; **19**: 1238-1249 [PMID: 20464432 DOI: [10.1007/s00586-010-1415-9](https://doi.org/10.1007/s00586-010-1415-9)]
- 24 **Audigé L**, Bhandari M, Hanson B, Kellam J. A concept for the validation of fracture classifications. *J Orthop Trauma* 2005; **19**: 401-406 [PMID: 16003200 DOI: [10.1097/01.bot.0000155310.04886.37](https://doi.org/10.1097/01.bot.0000155310.04886.37)]
- 25 **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**: 184-194 [PMID: 15180234 DOI: [10.1080/00016470412331294445](https://doi.org/10.1080/00016470412331294445)]

**P- Reviewer:** Emara KM, Yang RS

**S- Editor:** Dou Y **L- Editor:** A **E- Editor:** Tan WW





Published By Baishideng Publishing Group Inc  
7901 Stoneridge Drive, Suite 501, Pleasanton, CA 94588, USA  
Telephone: +1-925-2238242  
Fax: +1-925-2238243  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
Help Desk: <https://www.f6publishing.com/helpdesk>  
<https://www.wjgnet.com>

