

AO Spine-DGOU Osteoporotic Fracture Classification System: Internal Validation by the AO Spine Knowledge Forum Trauma

Global Spine Journal
2025, Vol. 15(4) 2152–2157
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/21925682241288187
journals.sagepub.com/home/gsj



Julian Scherer, MD^{1,2} , Andrei Joaquim, MD, PhD³ , Alex Vaccaro, MD, PhD⁴,
Rishi Kanna, MD⁵ , Mohammad El-Sharkawi, MD, PhD⁶ ,
Masahiko Takahata, MD, PhD⁷, Mohamed M. Aly, MD, PhD^{8,9} ,
Gaston Camino-Willhuber, MD¹⁰ , Ulrich Spiegl, MD¹¹ , Cumhur Oner, MD, PhD¹²,
Jose A. Canseco, MD, PhD⁴ , Ratko Yurac, MD¹³ , Lorin Michael Benneker, MD¹⁴,
Eugen Cezar Popescu, MD, PhD¹⁵ , Richard Bransford, MD¹⁶,
Harvinder Singh Chhabra, MD¹⁷, Frank Kandziora, MD, PhD¹⁸,
Marko H. Neva, MD, PhD¹⁹, and Klaus John Schnake, MD^{20,21}

¹ General Medicine & Global Health (GMGH), Department of Medicine and Orthopaedic Research Unit (ORU), Division of Orthopaedic Surgery, Faculty of Health Sciences, University of Cape Town, South Africa

² Department of Traumatology, University Hospital of Zurich, Zürich, Switzerland

³ Neurosurgery Division, Department of Neurology, State University of Campinas, Campinas-Sao Paulo, Brazil

⁴ Department of Orthopaedic Surgery, Rothman Institute, Thomas Jefferson University Hospital, Philadelphia, PA, USA

⁵ Spine Department of Orthopaedics and Spine Surgery, Ganga Hospital, Coimbatore, India

⁶ Department of Orthopaedic and Trauma Surgery, Faculty of Medicine, Assiut University, Assiut, Egypt

⁷ Department of Orthopaedic Surgery, Dokkyo Medical University School of Medicine, Soka, Japan

⁸ Department of Neurosurgery, Prince Mohammed Bin Abdulaziz Hospital, Riyadh, Saudi Arabi

⁹ Department of Neurosurgery, Mansoura University, Mansoura, Egypt

¹⁰ Department of Traumatology, Policina Gipuzkoa, San Sebastian, Spain

¹¹ Klinik für Unfallchirurgie und Orthopädie, Klinik München Harlaching, München, Germany

¹² Department of Orthopaedics, University Medical Centers, Utrecht, the Netherlands

¹³ Clinica Alemana de Santiago, University Del Desarrollo, Vitacura, Chile

¹⁴ Department for Spine Surgery, Sonnenhof Spital, University of Bern, Switzerland

¹⁵ Department of Neurosurgery, “Prof. N. Oblu” Emergency Hospital, Iasi, Romania

¹⁶ Department of Orthopaedics and Sports Medicine, Harborview Medical Center, University of Washington, Seattle, WA, USA

¹⁷ Sri Balaji Action Medical Institute, New Delhi, India

¹⁸ Center for Spine Surgery and Neurotraumatology, BG Unfallklinik Frankfurt Am Main, Frankfurt, Germany

¹⁹ Department of Orthopaedic and Trauma Surgery, Tampere University Hospital, Tampere, Finland

²⁰ Center for Spinal and Scoliosis Surgery, Malteser Waldkrankenhaus St. Marien Erlangen, Erlangen, Germany

²¹ Department of Orthopedics and Traumatology, Paracelsus Private Medical University Nuremberg, Nuremberg, Germany

Corresponding Author:

Julian Scherer, MD, General Medicine & Global Health (GMGH), Department of Medicine and Orthopaedic Research Unit (ORU), Division of Orthopaedic Surgery, Faculty of Health Science, University of Cape Town, Anzio Road, Chris Barnard Building, Observatory, Cape Town 7925, South Africa; Department of Traumatology, University Hospital Zurich, University of Zurich, Zurich, Switzerland.

Email: Julian.scherer@usz.ch



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work 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>).

Abstract

Study Design: Cross-sectional survey.

Objectives: Injury classifications are important tools to identify fracture patterns, guide treatment-decisions and aid to identify optimal treatment plans. The AO Spine-DGOU Osteoporotic Fracture (OF) classification system was developed, and the aim of this study was to assess the reliability of this new classification system.

Methods: 23 Members of the AO Spine Knowledge Forum Trauma participated in the validation process. Participants were asked to rate 33 cases according to the OF classification at 2 time points, 4 weeks apart (assessment 1 and 2). The kappa statistic (κ) was calculated to assess inter-observer reliability and intra-rater reproducibility. The gold master key for each case was determined by approval of at least 5 out of 7 members of the DGOU.

Results: A total of 1386 ratings (21 raters) were performed. The overall inter-rater agreement was moderate with a combined kappa statistic for the OF classification of 0.496 in assessment 1 and 0.482 in assessment 2. The combined percentage of correct ratings (compared to gold-standard) in assessment 1 was 71.4% and 67.4% in assessment 2. The average intra-rater reproducibility was substantial ($\kappa = 0.74$, median 0.76, range 0.55 to 1.00, SD 0.13) for the assessed fracture types.

Conclusions: The assessed overall inter-rater reliability was moderate and substantial in some instances. The average intra-rater reproducibility is substantial. It seems that appropriate training of the classification system can enhance inter- and intra-rater reliability.

Keywords

osteoporosis, spinal fracture, fracture classification, validation study

Background

Osteoporosis represents a major public health concern with the elderly population steadily rising worldwide.¹ Vertebral fractures are the most common pathological fractures as a result of low bone mass and bone fragility due to osteoporosis^{2,3} leading to morbidity and potential mortality.⁴ The management of osteoporotic vertebral fractures (OVFs) can be challenging due to patients' age, existing comorbidities, reduced physiological reserves and often multipharmacy.⁵ Osteoporotic changes in the elderly can either be a contributing factor in traumatic (low-energy) vertebral fractures or can be the cause of those fractures.⁴ Injury classifications are important tools for patient care as well as for communication amongst caregivers. Classification systems, ultimately, should guide physicians in decision-making for optimal treatment of the patient. Therefore, classification systems should be simple, reproducible and should accentuate important fracture characteristics which are relevant for patient care. Commonly used and internationally accepted classifications for (thoracolumbar) vertebral fractures were not specifically developed for patients with osteoporotic changes. This differentiation is important since OVFs behave and are managed differently compared to traumatic (high-energy) vertebral fractures.⁶

Recently, the Spine Section of the German Society for Orthopaedics and Trauma (DGOU), who have formed the working group "Osteoporotic Fractures", have proposed a new classification for osteoporotic thoraco-lumbar spine fractures.⁴

The final AO Spine-DGOU Osteoporotic Fracture (OF) classification system was developed after several consensus meetings with members the AO Spine Knowledge Forum Trauma. Blatter et al proposed treatment recommendations based on the OF classification: OF 1 and 2 should be treated

conservatively, whereas OF 3 can either be managed conservatively or surgically. OF 4 and OF 5 are unstable fractures which should be treated surgically.⁷ A large multi-center study from Germany, assessing performed treatment strategies of OVFs, showed a high accordance with the OF-based treatment recommendations.⁶ Whereas treatment recommendations were assessed and validated with existing clinical data, data on the reliability of the AO Spine-DGOU OF classification is lacking.

Thus, the aim of the present study was to assess the reliability of the AO Spine-DGOU OF classification amongst members of the AO Spine Knowledge Forum Trauma.

Material and Methods

Development of Classification System

The methodology behind the creation of the AO Spine-DGOU classification has been described in detail in the publication of the Spine Section of the German Society for Orthopaedics and Trauma (DGOU).⁴

The AO Spine-DGOU Osteoporotic Fracture (OF) Classification

This classification consists of 5 groups: OF 1, no deformation, only bone edema assessed with MRI, OF 2, deformation of 1 endplate with no or minor involvement of posterior wall, OF 3, deformation of 1 endplate with substantial involvement of posterior wall, OF 4, deformation of both endplates with or without involvement of posterior wall, and OF 5, fracture with anterior or posterior tension band failure. ([Supplement 1](#))

Validation Process

Members of the AO Spine Knowledge Forum Trauma, an international group of spinal surgeons dedicated to the study of spinal trauma from different regions of the world, participated in the validation process. Participants were asked to rate cases according to the OF classification at 2 time points, 4 weeks apart (assessment 1 and 2), to avoid recall bias. Participants were not particularly familiar with the classification but an introduction to the classification with example cases was given and participating surgeons were able to use the classification poster for the assessments and were asked to provide any feedback they may have to improve the classification and validation process. All of the selected cases included osteoporotic thoracolumbar fractures with different morphological subtypes. A total of 33 cases were selected, and 7 DGOU members developed the gold-standard master key for each case. In cases of disagreement, consensus was reached. All cases were reviewed independently by 23 surgeons from the AO Spine Knowledge Forum Trauma representing different regions of the world. To support the surgeons in rating the cases, key-images showing the fracture, and short video sequences of computed tomography and/or MRI were provided. Cases were randomized in both assessments.

Statistical Analysis

The kappa statistic (κ) was calculated to assess the reliability of the classification system among different observers (inter-rater

agreement) and the reproducibility for the same observer on separate occasions (intra-rater reproducibility).⁸ The Fleiss kappa statistic measures the agreement of multiple raters who rate multiple subjects, with the rating based on multiple categories. The coefficients were interpreted utilizing the Landis and Koch grading system, which defines $\kappa \leq 0.2$ as slight reliability (agreement/reproducibility), $0.2 < \kappa \leq 0.4$ as fair reliability, $0.4 < \kappa \leq 0.6$ as moderate reliability, $0.6 < \kappa \leq 0.8$ as substantial reliability, and $\kappa > 0.8$ as excellent reliability.⁹ Inter-rater and intra-rater agreement was calculated for the different classification groups (OF 1-5) and overall, for the classification itself. Furthermore, agreement with the gold-standard (accuracy) was calculated overall and for each OF-subtype.

Development of the OF Classification System

The classification consists of 5 different subtypes (OF 1-5)⁴ (Figure 1).

Results

Out of 23 raters, 2 were excluded from final analysis due to non-compliance in the assessment process. A total of 1386 (693 twice) ratings were performed. Of the 33 presented cases, 1 case was classified (gold-standard) as OF 1 (3%), ten as OF 2 (30.3%), 5 as OF 3 (15.1%), eleven as OF 4 (33.3%) and 6 as OF 5 (18.2%).

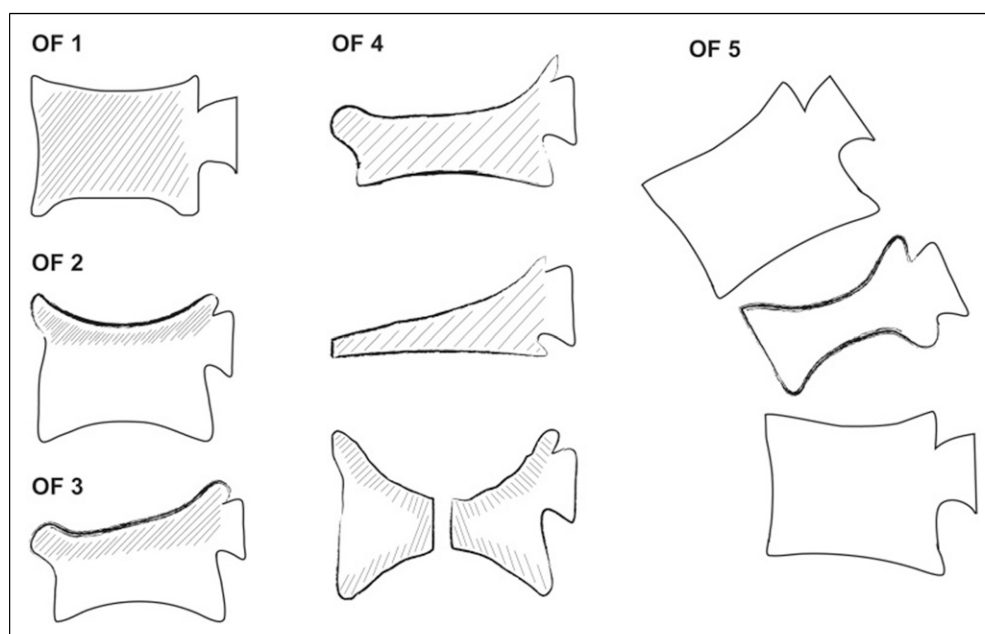


Figure 1. Development of the OF classification system. OF 1: No vertebral deformation (vertebral body edema in MRI-STIR only). X-rays do not show vertebral deformation. OF 2: Deformation of 1 endplate without or with only minor posterior wall involvement ($<1/5$). OF 3: Deformation of 1 endplate with distinct posterior wall involvement ($>1/5$). OF 4: Deformation of both endplates with or without posterior wall involvement. This type can lead to the loss of integrity of the vertebral frame structure (complete burst), or vertebral body collapse, or pincer-type fracture. OF 5: Injuries with anterior or posterior tension band failure. These injuries have signs of distraction, rotation, or translation.

Inter-Rater Agreement

The overall inter-rater agreement was moderate with a combined kappa statistic for the OF classification of 0.496 in assessment 1 and 0.482 in assessment 2, respectively. The lowest level of agreement was seen in OF 1 ($\kappa = 0.293$, 0.181 in assessment 2), whereas the highest level of agreement was observed in OF 4 ($\kappa = 0.625$, 0.683 respectively). (Table 1) In both assessments, the inter-rater levels of agreement for OF 2, 4 and 5 were higher than the overall level of agreement.

Agreement with Gold-Standard

The combined percentage of correct ratings (compared to gold-standard) in assessment 1 was 71.4% and 67.4% in assessment 2. (Table 2) In both assessments, the lowest level of agreement with the gold standard was seen in OF 1 and OF 5 fractures (61.9% and 38.1%, 59.5% and 67.5% respectively). Half of the raters assessed the OF 1 (1 case in each assessment) falsely as an OF 2. One of the 6 as OF 5 classified cases were rated wrongly by most of the raters (59.5%). The overall mean agreement with the gold-standard among raters in this study was 75.04% (median 75.76%, range 63.64% to 84.85%, SD 6.28) in assessment 1 and 75.90% (median 78.79%, range 57.58% to 84.85%, SD 8.65).

Intra-Rater Agreement

The average intra-rater reproducibility was substantial ($\kappa = 0.74$, median 0.76, range 0.55 to 1.00, SD 0.13) for the assessed fracture types. Three of the reviewers showed a moderate level of agreement ($0.4 < \kappa \leq 0.6$), whereas twelve raters had a substantial level of agreement ($0.6 < \kappa \leq 0.8$) and 6 of the reviewers were assessed with an excellent level of agreement ($\kappa > 0.8$).

Table 1. Inter-rater Agreement of the OF Classification.

Classification	Assessment 1 (κ)	Assessment 2 (κ)
Total	0.496	0.482
OF 1	0.293	0.181
OF 2	0.576	0.579
OF 3	0.469	0.363
OF 4	0.625	0.683
OF 5	0.518	0.603

Table 2. Agreement With Gold-Standard.

Classification	Assessment 1 (%)	Assessment 2 (%)
Combined	71.4	67.4
OF 1	61.9	38.1
OF 2	76.2	79.0
OF 3	76.2	67.6
OF 4	83.1	84.8
OF 5	59.5	67.5

Discussion

The prevalence of OVFs varies between approximately 19% in women and 20% in men amongst the elderly population.¹⁰ Due to a rising number of elderly people worldwide, the incidence rates of these fractures can be assumed to rise. OVFs represent a major public health concern due to increased morbidity and mortality, especially in patients with an age of 75 years and above.¹¹ Accurate diagnosis and treatment of these fractures is crucial to avoid prolonged and chronic episodes of pain, to reduce morbidity and mortality and to prevent kyphosis and further collapse of the fractured vertebra which can result in neurological deficits. However, to date, there is no consensus guideline on the treatment of OVFs and conservative treatment is considered the primary treatment option.¹² Existing classification systems on thoracolumbar vertebral fractures such as the AO Spine thoracolumbar spine injury classification system or the TLICS are widely accepted but are unsuitable for OVFs since these systems are not appreciating the distinct features of vertebral fractures with underlying osteoporotic changes.^{13,14} Recently, a new classification specifically addressing OVFs was introduced by the working group on "Osteoporotic Fractures" of the DGOU and was then further developed in cooperation with AO Spine (AO Spine-DGOU Osteoporotic Fracture (OF) classification system) which grades OVFs into 5 groups and gives treatment recommendations (surgical vs conservative) for each classification group.⁴ Thus, the aim of this study was to assess the inter-rater as well as intra-rater reliability amongst AO Spine Knowledge Forum Trauma members.

We assessed a moderate combined overall inter-rater agreement in both assessments ($\kappa = 0.496$ in assessment 1 and $\kappa = 0.482$ in assessment 2) which is in keeping with 2 previous studies assessing the inter-rater agreement of this classification amongst 6 surgeons in each study.^{15,16} Furthermore, we found that our findings are comparable to the previous international validation study of the AO Spine thoracolumbar injury classification system ($\kappa = 0.56$).¹⁷ Due to global osteoporotic changes of the vertebrae on the assessed imaging data, rating of the fractures might be more difficult (eg, degenerative endplate changes vs traumatic endplate changes) and therefore, the assessed inter-rater agreement might be falsely low. Interestingly, the inter-rater agreement declined between the first and the second assessment. OF 1 showed fair agreement ($\kappa = 0.293$) in assessment 1 and slight agreement ($\kappa = 0.181$) which represents the most inferior inter-rater agreement amongst the 5 classification groups. This relative low agreement rate may be also explained by the low prevalence of this morphology in the case sample as well as the difficulty to distinguish degenerative endplate irregularities from traumatic changes. OF 4 and OF 5 are recommended for surgical treatment and showed a substantial inter-rater agreement in assessment 2, which highlights the uniformity of rating and possibly treatment decisions.⁷ The assessed moderate overall inter-rater

agreement could be further due to lack of training of the included raters, which were all experienced raters with good training in well-established classifications who did not undergo training of the AO Spine-DGOU Osteoporotic Fracture (OF) classification system prior this assessment, which highlights the need for appropriate training of classification systems. This should be assessed further after appropriate training sessions of the respected classification system. We observed a combined percentage of correct ratings, compared to gold-standard, of 71.4% in assessment 1 and 67.4% in assessment 2. Like the inter-rater agreement, OF 1 was rated incorrectly the most often in both assessments and was rated worse in assessment 2 (61.9% and 38.1%). Half of the raters rated OF 1 as OF 2. This could be due to over-interpretation of MRI imaging ie, interpreting physiological endplate deformities with bone marrow edema as impression fracture. Therefore, assessment of both, CT and MRI are necessary to avoid an over-interpretation and to increase inter-rater and agreement with the gold-standard, which has been also shown in a previous assessment of 6 surgeons.^{16,18} According to the newly introduced OF-score, both fracture types, in the worst clinical scenario, could be treated surgically.⁷ However, a multi-center study showed that OF 1 are almost always and OF 2 are commonly treated conservatively.⁶ A possible second solution to increase the intra-rater agreement as well as the agreement with the gold-standard would therefore be to fuse OF 1 and OF 2 into 1 entity and make surgical decision-making solely dependent on the modifiers according to the OF-score.

We assessed a substantial overall intra-rater reproducibility of 0.74 which is comparable to the AO Spine thoracolumbar spine injury classification system ($\kappa = 0.68$). This finding is also comparable to the assessment from Germany, which found a substantial intra-observer agreement when raters were provided with radiograph and MRI ($\kappa = 0.64$). This study, in addition, showed that the intra-observer reliability can be increased by providing CT-scans as an additional imaging modality, which highlights the importance of complete imaging.^{16,17} The majority of the assessed raters showed a substantial or excellent reproducibility.

This study has certain limitations. Firstly, there might be a varying degree of familiarity with the OF-classification amongst the raters which results in a classification knowledge-inequality and could result in a potential bias. Furthermore, the sole use of key-images (in some cases) or vice versa might have an impact on the rater's response. It is likely that some of the reviewers would change their response while assessing a complete imaging file rather than solely key images. Another potential limitation is the unequal frequency of different fracture subtypes and hence, the data on rare subtypes (eg, OF 1) should be interpreted with caution. Additionally, some of the presented images showed multiple fractures which might have caused confusion regarding the fracture to rate, decreasing the reliability rate.

Furthermore, the members of the AO Spine Knowledge Forum Trauma might have varying degrees of knowledge on interpreting different imaging modalities and familiarity with the presented classification since not all members were involved in the development process, which could result in an inhomogeneous group of raters and could have a negative effect on the inter-rater agreement. Standardized training on imaging and of the OF-classification could limit these potential biases.

Conclusion

Like previously described AO Spine thoracolumbar fracture classifications as well as small inter- and intra-rater reliability assessments, the results from the present study showed moderate to substantial interrater reliability and substantial inter-rater reproducibility amongst a group of members of the AO Spine Knowledge Forum Trauma. As a next step, an international validation amongst spine surgeons with a higher number of raters representing different backgrounds, work settings, and regions should be performed to assess the AO Spine-DGOU OF classification reliability globally. It seems that appropriate training of the respected classification system is necessary to increase inter- and intra-rater reliability.

Acknowledgments

The authors of the manuscript would like to thank Olesja Hazenbiller for her assistance in developing the methodology and providing support during the validation. We would also like to thank Dr Stefan Bentink Director Data Science and Diagnostics from Staburo GmbH for his assistance with the statistical analysis. We also would like to extend our gratitude to the Osteoporotic Fracture working group of the German Society for Orthopaedics and Trauma (DGOU) for their first draft of the classification system.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The AO Spine-DGOU Injury Classification System was developed and funded by AO Spine through the AO Spine Knowledge Forum Trauma, a focused group of international spine trauma experts. AO Spine is a clinical division of the AO Foundation, which is an independent medically-guided not-for-profit organization. Individual authors of this manuscript have no disclosures to declare regarding the proposed article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: 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 AO Network Clinical Research.

ORCID iDs

Julian Scherer  <https://orcid.org/0000-0001-8943-5817>
 Andrei Joaquim  <https://orcid.org/0000-0003-2645-0483>
 Rishi Kanna  <https://orcid.org/0000-0001-5817-4909>
 Mohammad El-Sharkawi  <https://orcid.org/0000-0001-6177-7145>
 Mohamed M. Aly  <https://orcid.org/0000-0001-5221-8288>
 Gaston Camino-Willhuber  <https://orcid.org/0000-0002-5684-7679>
 Ulrich Spiegl  <https://orcid.org/0000-0002-5179-4192>
 Jose A. Canseco  <https://orcid.org/0000-0002-2152-5725>
 Ratko Yurac  <https://orcid.org/0000-0003-3603-6294>
 Eugen Cezar Popescu  <https://orcid.org/0000-0001-5732-5402>

Data Availability Statement

Data is available upon reasonable request towards the corresponding author.

Supplemental Material

Supplemental material for this article is available at: <https://www.aofoundation.org/spine/clinical-library-and-tools/aospine-classification-systems>

References

1. Reginster JY, Burlet N. Osteoporosis: a still increasing prevalence. *Bone*. 2006;38(2 Suppl 1):S4-9.
2. Ensrud KE. Epidemiology of fracture risk with advancing age. *J Gerontol A Biol Sci Med Sci*. 2013;68(10):1236-1242.
3. Khavinson V, Popovich I, Mikhailova O. Towards realization of longer life. *Acta Biomed*. 2020;91(3):e2020054.
4. Schnake KJ, Blattert TR, Hahn P, et al. Classification of osteoporotic thoracolumbar spine fractures: recommendations of the spine section of the German society for Orthopaedics and trauma (DGOU). *Global Spine J*. 2018;8(2 Suppl):46s-49s.
5. Hadji P, Klein S, Gothe H, et al. The epidemiology of osteoporosis—Bone Evaluation Study (BEST): an analysis of routine health insurance data. *Dtsch Arztebl Int*. 2013;110(4):52-57.
6. Ullrich BW, Schenk P, Scheyerer MJ, et al. Georg Schmorl prize of the German spine society (DWG) 2022: current treatment for inpatients with osteoporotic thoracolumbar fractures—results of the EOFTT study. *Eur Spine J*. 2023;32(5):1525-1535.
7. Blattert TR, Schnake KJ, Gonschorek O, et al. Nonsurgical and surgical management of osteoporotic vertebral body fractures: recommendations of the spine section of the German Society for Orthopaedics and Trauma (DGOU). *Global Spine J*. 2018;8(-2_suppl):50S-55S.
8. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull*. 1971;76(5):378-382.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
10. Waterloo S, Ahmed LA, Center JR, et al. Prevalence of vertebral fractures in women and men in the population-based Tromsø Study. *BMC Musculoskel Disord*. 2012;13:3.
11. Cauley JA. Public health impact of osteoporosis. *J Gerontol A Biol Sci Med Sci*. 2013;68(10):1243-1251.
12. Jang HD, Kim EH, Lee JC, Choi SW, Kim K, Shin BJ. Current concepts in the management of osteoporotic vertebral fractures: a narrative review. *Asian Spine J*. 2020;14(6):898-909.
13. Vaccaro AR, Oner C, Kepler CK, et al. AOSpine thoracolumbar spine injury classification system: fracture description, neurological status, and key modifiers. *Spine*. 2013;38(23):2028-2037.
14. Vaccaro AR, Lehman RA Jr., Hurlbert RJ, et al. A new classification of thoracolumbar injuries: the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. *Spine*. 2005;30(20):2325-2333.
15. Quinteros G, Cabrera JP, Urrutia J, et al. Reliability evaluation of the new AO spine-DGOU classification for osteoporotic thoracolumbar fractures. *World Neurosurg*. 2022;161:e436-e440.
16. Schonrogge M, Lahodski V, Otto R, et al. Inter- and intra-observer reliabilities and critical analysis of the osteoporotic fracture classification of osteoporotic vertebral body fractures. *Eur Spine J*. 2022;31(9):2431-2438.
17. 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(4):1082-1086.
18. Spiegl UJ, Behr L, Osterhoff G, Rupprecht G, Scheyerer MJ, Katscher S. OF spine classification of osteoporotic thoracolumbar vertebral body fractures by MRI and conventional radiographs only leads to high inter-observer agreement rates—an additional CT adds limited information for the of classification and the OF score. *BMC Musculoskel Disord*. 2022;23(1):1086.