SO KATO, MD, PhD

So Kato, MD, PhD, is an Assistant Professor in the Department of Orthopaedic Surgery at the University of Tokyo, Japan. Dr Kato pursued his advanced fellowship in complex spine surgery at the University of Toronto, Canada, where he served as a chief clinical fellow. He is a board-certified orthopaedic surgeon who specializes in spinal metastasis, spinal deformity, and degenerative cervical myelopathy.

doi: 10.6004/jnccn.2020.7544

The ideas and viewpoints expressed in this commentary are those of the author and do not necessarily represent any policy, position, or program of NCCN.

Developing a Reliable and Versatile System for the Diagnosis of Metastatic Vertebral Fracture: Challenges and Future Directions

So Kato, MD, PhD

Distinguishing metastatic vertebral fractures (MVF) from osteoporotic vertebral fractures (OVFs) has long been a challenge among healthcare providers, but the clinical significance of this difference has increased in light of the increasing number of cancer survivors and geriatric osteoporosis patients encountered in daily practice. Indeed, the populations among which these 2 diagnoses are commonly made have significant overlap. A delay in making a correct diagnosis of metastatic disease can have devastating consequences.

The biomechanical backgrounds of vertebral fragility that lead to fracture and collapse differ between MVF and OVF: MVF occurs when spotty osteolysis, in which bone trabeculae are replaced by tumor tissue, grows large enough to impair structural integrity; OVF is caused by universal trabecular thinning that renders a vertebra unable to tolerate the physiologic load.

MRI is considered the imaging modality with the strongest discriminant capability, enabling visualization of the structural changes and determination of the typical patterns of collapse, depending on the etiology.¹ For example, a band-like signal intensity change in a fractured vertebra on sagittal images implies intramedullary hemorrhaging due to axial loading evenly distributed from the anterior to the middle column. In contrast, MVF is typically associated with the concentric expansion of the osteolytic area rich in cellular components. Likewise, the signal changes that involve posterior elements (ie, pedicles and laminae) are unlikely to be seen in typical cases of OVF due to the compressive force applied to the spinal axis. The morphology of posterior wall protrusion also aids in the differentiation of MVF from OVF. A diffusely convex border representing the soft consistency of the protruded wall indicates MVF, whereas focal retropulsion with a “sharp” edge indicates the bony spike seen in OVF. In addition, contour in axial images is well recognized as a typical radiologic sign that supports the diagnosis of malignancy. Spinal canal encroachment seen in MVF often has 2 peaks, with the reason believed to be the intact posterior longitudinal ligament functioning as a strong physical barrier in the center.

However, none of these key features has perfect sensitivity or specificity. Therefore, no single finding by itself is pathognomonic, and the lack of or existence of these signs must be treated with careful consideration. Such consideration demands advanced clinical experience, through occasionally lengthy training periods.

Scoring systems have been designed to simulate these experts’ integration process.^{2–4} The virtue of these scoring systems is that they do not require their users to intuitively determine the clinical relevance of each aspect of a radiologic image. For example, if “finding A,” which supports MVF, exists in the patient’s MRI scan, but “finding B,” which strongly negates its possibility, also exists at the same time, the fracture can be concluded to be OVF. In other words, the rater needs to understand the relative importance of different findings and to integrate them after proper weighting. In the scoring system, this process is automatically structured by assigning scores to the observed findings used.

See page 267 for related article.

Colleagues and I published on the MRI Evaluation Totalizing Assessment (META) score in 2015.³ This system uses only MRI features to determine the score. In that study, 7 MRI findings that were thoroughly reported in the literature were analyzed across 100 OVFs and 100 MVFs to evaluate sensitivity and specificity. Using these findings as variables, a discriminant analysis was performed using 140 fractures as a training set. The classification accuracy was verified using another 60 fractures as a test set. This score gained popularity as a fundamental screening tool. As shown in Table 1, some findings have greater relevance than others for drawing conclusions.

However, approaches such as the META score are practically valid based on the assumption of acceptably reliable radiologic interpretation. The META score's reproducibility has been questioned, particularly among multidisciplinary healthcare providers.^{5,6}

In this issue, Arana et al⁷ highlight this debate. In their well-designed study, imaging findings from 203 patients with confirmed diagnoses of MVF or OVF were provided to 25 clinicians, and the interrater and intrarater reliability of the key features that discriminate MVF from OVF, including those included in the META score, were tested, along with the concordance between the clinical diagnosis and biopsy or long-term follow-up results. The intrarater reliability of the MRI findings was "moderate" to "substantial," but the inter-rater reliability was only "fair" to "moderate." The diagnostic accuracy was also only "moderate"; kappa value was 0.452 without preexisting fractures, which only improved to 0.462 after disclosing patients' clinical history of cancer. The authors concluded that the diagnostic power of MRI for MVF was questionable at best.

Given that MRI remains the mainstay modality for diagnosing MVF and OVF, these results are disappointing. However, they must be interpreted with caution. One of the challenges associated with the study design is that when the findings are interpreted as binary parameters ("yes" or "no"), their results depend on a threshold or definition arbitrarily set by each rater. Thus, a significant amount of information is lost in the process. Furthermore, as the authors mention, the entire series of the cutting surfaces is reviewed in real-world practice, not only certain key slides; this enables physicians to grasp the 3-dimensional aspect to further characterize the fractures. Because the experts' diagnoses are made considering the nuance of these constellations of subtle signs, the marginal reliability of each finding does not necessarily deny the modality's usefulness in the clinical setting.

Now that we have entered the era of artificial intelligence, our hope is that a more systematic method for the differential diagnosis of MVF and OVF will be established through deep learning. Chea and Mandell⁸ showcased state-of-the-art examples of the recent application of this technology in musculoskeletal radiology. Deep learning with convolutional neural networks is extremely useful for lesion detection. In the classic form of machine learning, human input by experts is required to determine which imaging features are most important in the

Table 1. META Score

MRI Findings	Score
Sagittal images	
Shape of vertebral signal intensity change	
Diffuse	6
Round	7
Band-like	-5
Other	0
Cleft formation	
Yes	-3
No	0
Signal intensity change in the pedicles	
Yes	3
No	0
Posterior wall protrusion	
Blunt	6
Sharp	-5
None	0
Axial images	
Asymmetric signal intensity change	
Yes	5
No	0
Total score of ≤ 4 indicates OVF; ≥ 5 indicates MVF.	

Data from Kato S, Hozumi T, Yamakawa K, et al. META: an MRI-based scoring system differentiating metastatic from osteoporotic vertebral fractures. *Spine J* 2015;15:1563-1570.

diagnosis to create the algorithm, just as we have attempted in the development of our scoring system. However, deep learning allows the system itself to determine the imaging features that characterize a certain condition. Even experienced radiologists may not be aware of all of these features. Therefore, the reliability or reproducibility of human radiologic interpretation may no longer be an issue, because the same program will be used every time and the diagnostic procedure will automatically apply systematic integration through neural networks. Ideally, an automated system that checks all MR images to screen for vertebral fractures and sends an alert when MVF is highly suspected should be introduced.

However, it should be noted that diagnostic tools, no matter how sophisticated, are useless without the suspicion of malignancy from clinicians. Unless all imaging studies are evaluated using an automatic diagnostic algorithm—which may not be realistic any time soon—the first action is always started by human physicians encountering patients with back pain who have red flags or a known history of malignancy. Even if the correct diagnosis of MVF cannot be made after the first MRI scan for back pain, follow-up MRI a few months later is advised. The worst approach to take is to be unaware of or indifferent in such cases.

Disclosures: The author has disclosed that he has no financial interests, arrangements, or affiliations with the manufacturers of any products discussed in this article or their competitors.

Correspondence: So Kato, MD, PhD, Department of Orthopaedic Surgery, University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-8655, Japan. Email: skatou-tyk@umin.org

References

1. Thawait SK, Marcus MA, Morrison WB, et al. Research synthesis: what is the diagnostic performance of magnetic resonance imaging to discriminate benign from malignant vertebral compression fractures? Systematic review and meta-analysis. *Spine (Phila Pa 1976)* 2012;37:E736–744.
2. Yuzawa Y, Ebara S, Kamimura M, et al. Magnetic resonance and computed tomography-based scoring system for the differential diagnosis of vertebral fractures caused by osteoporosis and malignant tumors. *J Orthop Sci* 2005;10:345–352.
3. Kato S, Hozumi T, Yamakawa K, et al. META: an MRI-based scoring system differentiating metastatic from osteoporotic vertebral fractures. *Spine J* 2015;15:1563–1570.
4. Li Z, Guan M, Sun D, et al. A novel MRI- and CT-based scoring system to differentiate malignant from osteoporotic vertebral fractures in Chinese patients. *BMC Musculoskelet Disord* 2018;19:406.
5. Urrutia J, Besa P, Morales S, et al. Does the META score evaluating osteoporotic and metastatic vertebral fractures have enough agreement to be used by orthopaedic surgeons with different levels of training? *Eur Spine J* 2018;27:2577–2583.
6. Besa P, Urrutia J, Campos M, et al. The META score for differentiating metastatic from osteoporotic vertebral fractures: an independent agreement assessment. *Spine J* 2018;18:2074–2080.
7. Arana E. Metastatic versus osteoporotic vertebral fractures on MRI: a blinded, multicenter, and multispecialty observer agreement evaluation. *J Natl Compr Canc Netw* 2020;18:267–273.
8. Chea P, Mandell JC. Current applications and future directions of deep learning in musculoskeletal radiology. *Skeletal Radiol* 2020;49:183–197.

Recorded Presentations

NCCN 2019 Annual Congress:

Hematologic Malignancies™

Earn CME/CNE/CPE credit online with recordings from the September 2019 Congress held in San Francisco.

NCCN 2019 Annual Congress: Hematologic Malignancies™

Earn credits online!

Interactive sessions include debates, panel discussions on patient case studies, and updates to the standard of care as recommended in the NCCN Guidelines®.

Visit education.nccn.org/hem2019

NCCN Nursing Forum: Advancing Oncology Nursing in Hematologic Malignancies™

Earn contact hours online!

Provides hematology/oncology nurses with practical information that can be implemented in the practice setting to optimize patient education and care.

Visit education.nccn.org/forum2019

JNCCN-N-0281-0320