

Original Article

Received: September 20, 2017
Revised: November 8, 2017
Accepted: November 28, 2017

Correspondence to:

Joon Woo Lee, M.D.
Department of Radiology, Seoul
National University Bundang
Hospital, 300 Gumi-dong,
Bundang-gu, Seongnam 13620,
Korea.
Tel. +82-31-787-7609
Fax. +82-31-787-4011
E-mail: joonwoo2@gmail.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2018 Korean Society of Magnetic Resonance in Medicine (KSMRM)

Do Radiology Residents Perform Well in Preliminary Reporting of Emergency MRIs of Spine?

Joon Woo Lee^{1*}, Guen Young Lee^{1,2}, Le Roy CHONG³, Heung Sik Kang¹

¹Department of Radiology, Seoul National University Bundang Hospital, Seongnam, Korea

²Department of Radiology, Chung-Ang University Hospital, Seoul, Korea

³Department of Radiology, Changi General Hospital, Singapore, Singapore

Purpose: To evaluate interpretation errors involving spine MRIs by residents in their second to fourth year of training, classified as minor, intermediate and major discrepancies, as well as the types of commonly discordant lesions with or without clinical significance.

Materials and Methods: A staff radiologist evaluated both preliminary and final reports of 582 spine MRIs performed in the emergency room from March 2011 to February 2013, involving (1) the incidence of report discrepancy, classified as minor if there was sufficient description of the main MR findings without ancillary or incidental lesions not influencing the main diagnosis, treatment, or patients' clinical course; intermediate if the correct diagnosis was made with insufficient or inadequate explanation, potentially influencing treatment or clinical course; and major if the discrepancy affected the main diagnosis; and (2) the common causes of discrepancy. We analyzed the differences in the incidence of discrepancy with respect to the training years of residents, age and sex of patients.

Results: Interpretation discrepancy occurred in 229 of the 582 cases (229/582, 39.3%), including 146 minor (146/582, 25.1%), 40 intermediate (40/582, 6.9%), and 43 major cases (43/582, 7.4%). The common causes of major discrepancy were: over-diagnosis of fracture (n = 10), missed cord lesion (n = 9), missed signal abnormalities associated with diffuse marrow (n = 5), and failure to provide differential diagnosis of focal abnormal marrow signal intensity (n = 5). No significant difference was found in the incidence of minor, intermediate, and major discrepancies according to the levels of residency, patients' age or sex.

Conclusion: A 7.4% rate of major discrepancies was found in preliminary reporting of emergency MRIs of spine interpreted by radiology residents, probably related to a relative lack of clinical experience, indicating the need for additional training, especially involving spine trauma, spinal cord and bone marrow lesions.

Keywords: Spine; Magnetic resonance imaging; Emergency; Resident; Discrepancy

INTRODUCTION

Although many staff radiologists try to teach and show key illustrative cases to their residents during the short four-year residency period, a few residents report inadequate training curriculum, especially in the area of spine MRI (1). Previous studies

investigated the discrepancies between preliminary reports by residents and final study results by staff radiologists (2–8). These included neuroradiology and spine studies read across all levels of residency, focusing on the accuracy of preliminary reading based on their clinical impact or factors affecting interpretation errors (6–8). However, it may be essential to disclose the types of common interpretation errors by radiologists for the benefit of residents with at least 1 year of training, who may be tasked with the reading and interpretation of spine MRI (9). Our hospital has a dedicated spine radiology section, and to the best of our knowledge, there has been no study evaluating the adequacy of preliminary reports by radiology residents and the common types of discrepancy, especially involving spine MRIs. We hypothesize that major interpretation errors of spine MRI reported by residents may be similar to previous studies involving other subspecialties such as neuroimaging (6–8). Therefore, this retrospective study evaluated the interpretation errors involving spine MRIs made by residents in their second to fourth year of training, classified as minor, intermediate and major discrepancies, as well as the different types of common discordant lesions with or without clinical significance.

MATERIALS AND METHODS

This retrospective study was approved by the Institutional Review Board of our hospital, and informed consent was waived.

Patients

From March 2011 to February 2013, 600 spine MRIs were performed on patients in the emergency room. The MRIs included 582 cases with preliminary reports generated by radiology residents. They included 288 males (mean age, 52.7 years; range of age, 3–89 years) and 294 females (mean age, 55.5 years; range of age, 16–92 years) with a mean age of 54.1 years (range, 3–92 years).

Spine MRI

The main indications for spine MRI were: pain ($n = 275$), trauma ($n = 184$), cancer evaluation ($n = 41$), weakness ($n = 40$), fever ($n = 36$), sensory change ($n = 4$), suspected postoperative complication ($n = 1$), and visual disturbance ($n = 1$). The study involved 151 cervical, 30 thoracic, 298 lumbosacral, and 103 whole spine MRIs including 329 with contrast enhancement. Two hundred eighty seven

lumbosacral MRI examinations included T2-weighted sagittal scans at the cervicothoracic level. Two hundred and seventy two cases were investigated using a 3T MRI scanner (Achieva, Philips Healthcare, the Netherlands) and the others with a 1.5T scanner (Intera, Philips Healthcare, the Netherlands). None of the MRIs was poor in image quality that precluded interpretation by the residents. Most of the emergency MRIs involving spine were reviewed by 4 second-year (284 exams) or 3 third-year (293 exams) radiology residents. Only four exams were reported preliminarily by 2 first-year residents and one by a fourth-year resident.

Image Analysis

Based on the clinical indication for emergency MRIs of spine, the main diagnosis related to patient's chief complaint. For example, if a patient complained about pain without trauma history, degenerative change, disc abnormalities, or tumorous lesion may represent a possible primary diagnosis. If a patient has trauma history, vertebral fracture or other soft tissue injuries may be a related diagnosis. If a patient shows weakness or sensory changes, myelopathy associated with stenosis or inflammatory myelitis may be a possible symptom. If a patient has fever, it may be essential to determine whether a finding indicative of spondylodiscitis or soft tissue infection exists. Interpretation errors involving preliminary reports by radiology resident and the final report by staff radiologists were evaluated, and the degree of discrepancy was classified into minor, intermediate, and major types. A minor discrepancy was defined as inadequate description of a patient's chief complaint even after the main finding was described sufficiently, but without mention of ancillary findings or incidental lesion, not influencing the main diagnosis, treatment decision, or patients' clinical course. An intermediate discrepancy was defined as one in which the main finding of spine MRI was identified without insufficient or inadequate explanation, and which may potentially influence treatment options or clinical course, without altering the main diagnosis. A major discrepancy was regarded as a severe error that potentially altered the main diagnosis. According to the degree of discrepancy, we also evaluated the types of commonly discordant lesions with or without clinical significance.

Statistical Analysis

We investigated the significant differences in the degree of discrepancy with respect to the number of years of training of the residents, age and sex of patients, using the

chi-square test or one-way analysis of variance (ANOVA). A P value of less than 0.05 was deemed as significant difference. To determine a difference between two resident groups by the number of training years, a P value of less than alpha value using Bonferroni post Hoc test was considered to indicate significant difference. Statistical analysis was performed using statistical software (PASW, version 17.0; SPSS, Chicago, IL, USA).

Table 1. Main Final Diagnosis of the 582 Emergency Spine MRIs

Main diagnosis related to patients' chief complaint	Number of spine MRI
HIVD	183
Fracture with/without soft tissue or cord injury	113
Spinal stenosis with/without compressive myelopathy	85
No acute traumatic lesion	62
Bone metastasis with/without pathologic fracture	33
Infectious spondylodiscitis	28
Normal	21
Non-tumorous myelopathy (ATM, MS, NMO, GBM)	15
Postoperative complications (e.g., postoperative abscess)	6
Leptomeningeal/intramedullary metastasis	5
No spinal metastasis	5
Spinal abscess	4
Cord contusion	3
Cord ischemia/infarction	3
Lymphoma (including CNS lymphoma)	3
Multiple myeloma	3
Sacral insufficiency fracture	2
Acute leukemia	1
Cellulitis	1
IDEM (intradural extramedullary tumor, neurogenic tumor)	1
Neurofibromatosis type 1	1
Neurofibromatosis type 2	1
Medullary infarction	1
Osteosarcoma	1
Spinal gout	1
Total	582

ATM = acute transverse myelitis; CNS = central nervous system; GBM = Gullain-Barre syndrome; HIVD = herniated intervertebral disc; IDEM = intradural extramedullary mass; MS = multiple sclerosis; NMO = neuromyelitis optica

RESULTS

The main diagnoses related to the patients' chief complaints are shown in Table 1. HIVD (herniation of intervertebral disc, $n = 183$) was the most common diagnosis and trauma-related lesions ($n = 113$) were the next most common, followed by spinal stenosis ($n = 85$).

Preliminary reports of more than half of the cases (353/582, 60.7%) were concordant with the final one. However, the important MR findings were discordant in 40 cases classified as intermediate discrepancies (40/582, 6.9%), while 43 were classified as major discrepancies, which changed the primary final diagnosis, following review by staff radiologists (43/582, 7.4%). These are listed in Table 2.

In the 43 cases with major discrepancies, over-diagnosis of fractures was the most common cause ($n = 10$), due to confusion with spondylolysis, and vascular grooves (Fig. 1). Abnormal spinal cord signaling suggesting compressive myelopathy or cord contusion ($n = 9$), was the second most frequently missed diagnosis. Diffuse abnormal marrow signal change ($n = 5$) also was missed by residents (Fig. 2).

Among indeterminate discrepancies, incomplete description of the main diagnosis, such as failure to discuss sequestration or migration of HIVD ($n = 12$), and soft tissue injuries including posterior ligamentous complex (PLC) or discoligamentous complex (DLC) disruption ($n = 11$, Fig. 3), were the leading factors underlying the errors.

For minor discrepancies, missed lesions at another level outside the main level of abnormality ($n = 47$) was the most frequent reason, other common reasons included failure to describe stenosis ($n = 26$) and missed spinal lesion outside the scanned levels of interest ($n = 34$).

There was no significant difference in the incidence of minor, intermediate, and major discrepancies with residents' years of training ($P = 0.066$, chi-square test). No difference was observed between second- and third year residents, in the incidence of minor, intermediate, and major discrepancies ($P = 0.024$, chi-square test; as significant if $P < \alpha$, 0.017 after Bonferroni post Hoc test), as well as incidence of no discrepancy, minor, intermediate, and major discrepancies ($P = 0.015$, chi-square test; as significant if $P < \alpha$, 0.0124 after Bonferroni post Hoc test). There was no difference in the incidence of minor, intermediate, and major discrepancies according to patients' age ($P = 0.489$, one-way ANOVA) or sex ($P = 0.161$, chi-square test).

No clinically adverse event related to the discrepant preliminary reports, as the final reports were issued within

Table 2. Causes of Interpretation Discrepancies

Degree of discrepancy	Number of spine MRIs
No discrepancy	353
Minor discrepancy	146
Missed lesion at level other than the level of the main finding	47
Missed lesion outside the scanned levels of interest (cervicothoracic spine, sacrum, coccyx, and other sites)	34
Missed foraminal stenosis	26
Benign bone marrow lesion, including hemangioma not mentioned	11
Ancillary finding associated with the main finding (epidural/subdural hemorrhage, pre/paravertebral hemorrhage)	7
OPLL or OLF not mentioned	7
Missed epidural lipomatosis	5
To distinguish postoperative status not mentioned	3
Failed in distinguishment between hemangioma and Schmorl's node	2
Missed fibrolipoma of filum terminale	1
IDEM (probable neurogenic tumor) not mentioned	1
Missed perineural cyst	1
Old spinal fracture not mentioned	1
Intermediate discrepancy	40
Incomplete description about HIVD (sequestration, migration, extraforaminal location, faulty nerve root)	12
Failed detection of soft tissue injury, PLC or DLC injury	11
Missed fracture of the posterior compartment	5
Failed to distinguish between compression and burst fractures	5
Over-diagnosis of soft tissue lesion, including PLC or DLC injury	2
Combined pathologic fracture with/without central canal compromise not mentioned	2
Missed combined leptomeningeal or intramedullary metastasis	2
Missed conjoined nerve root	1
Major discrepancy	43
Over-diagnosis of fracture, e.g. confusing with spondylolysis, vascular grooves, etc.	10
Missed abnormal spinal cord signal intensity (compressive myelopathy, cord contusion)	9
Missed abnormal diffuse marrow signal intensity	5
Failed to provide differential diagnose of focal abnormal marrow signal intensity between fracture, infection, metastasis and hemangioma	5
Missed HIVD, including HIVD recurrence	3
Failed in distinguishment between benign and malignant fractures	3
Failed to distinguish between acute and old fractures	3
Mistake in identifying level of the main spinal level (lumbarization, sacralization)	2
Missed acute soft tissue injury, including PLC or DLC injury	1
Mistake in identifying the main level of acute trauma/fracture	1
Missed acute spinal fracture	1
Total	582

DLC = discoligamentous complex; HIVD = herniated intervertebral disc; IDEM = intradural extramedullary mass; OLF = ossified ligamentum flavum; OPLL = ossified posterior longitudinal ligament; PLC = posterior ligamentous complex



Fig. 1. Lumbar spine MRI of a 39-year-old woman with back pain. Linear lesion at the spinous process of L3 presenting high signal intensity on T2-weighted mid-sagittal image (a, arrow) and low signal intensity on T1-weighted mid-sagittal image (b, arrow), which enhances the vascular structure along the spinous process on T1-weighted enhanced mid-sagittal (c, arrow) and axial scans (d, arrow). This lesion was described as a fracture of spinous process, and determined as a major discrepancy.

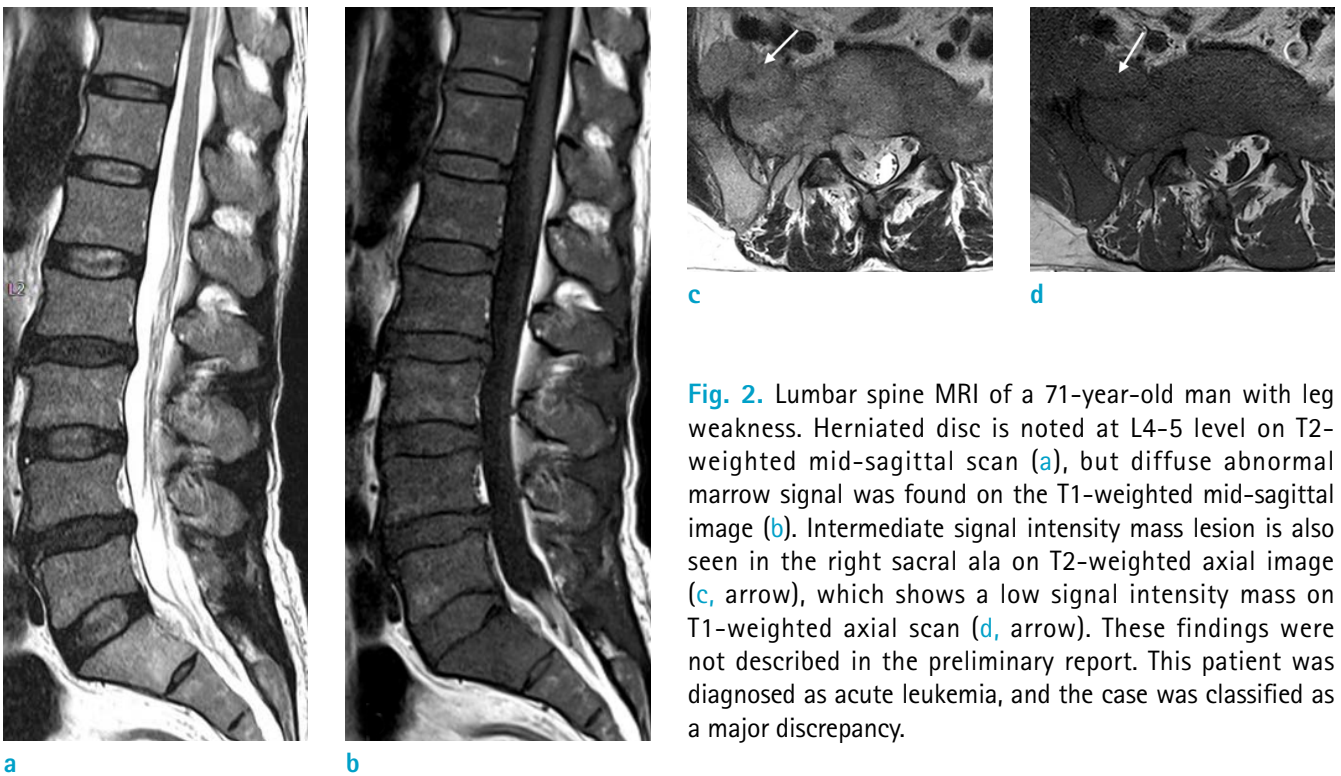


Fig. 2. Lumbar spine MRI of a 71-year-old man with leg weakness. Herniated disc is noted at L4-5 level on T2-weighted mid-sagittal scan (a), but diffuse abnormal marrow signal was found on the T1-weighted mid-sagittal image (b). Intermediate signal intensity mass lesion is also seen in the right sacral ala on T2-weighted axial image (c, arrow), which shows a low signal intensity mass on T1-weighted axial scan (d, arrow). These findings were not described in the preliminary report. This patient was diagnosed as acute leukemia, and the case was classified as a major discrepancy.

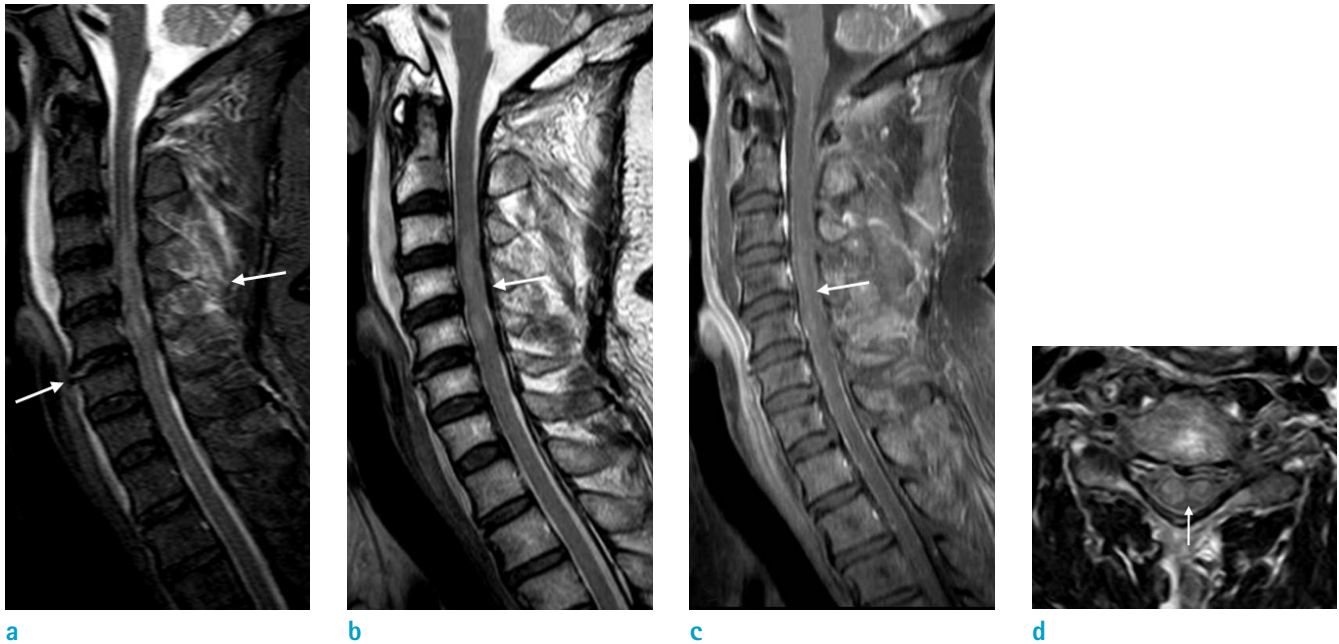


Fig. 3. Cervical spine MRI of a 66-year-old man following a traffic accident. Linear intra-discal high signal lesion with prevertebral hemorrhage and high signal intensity of the posterior ligamentous complex suggested traumatic discoligamentous injury on T2-weighted STIR sagittal image (a, arrows). Intramedullar high signal intensity with cord swelling is also observed on both T2-weighted STIR (a) and T2-weighted mid-sagittal scan (b, arrow), with enhancement on the T1-weighted enhanced mid-sagittal image (c, arrow), indicative of cord injury. This finding is also clearly visible on the T2-weighted axial image (d, arrow). The traumatic lesions including the acute cord injury were not mentioned in the preliminary report by the on-call resident, and were classified as a major discrepancy.

24 hr and results conveyed to clinicians. The clinicians usually commenced appropriate treatment based on patients' clinical features although the preliminary report may be non-diagnostic.

DISCUSSION

Before initiating this retrospective study, we searched Google for the distribution of distinguished spine radiology institutions using a keyword "dedicated spine radiology section." We found "American Society of Spine Radiology", which suggested that many countries may not have a dedicated spine imaging section in their hospital. Many trainees including fellows and residents feel that spine imaging is rather difficult and underscored the need for stronger clinical experience and more robust curriculum involving spinal imaging (1). This experience may be related in part to radiology departments in most countries lacking a dedicated spine imaging section, and preference of residents for other sub-specialties such as chest or

abdominal imaging. Our hospital has had its own dedicated spine radiology section for the past 10 years, and all the residents rotate through the spine radiology section for a month each year. All the residents from the second year of training are also responsible for overnight on-call duty to attend to emergency and urgent cases. In our hospital, 2 to 3 cases of spine MRI usually were read primarily by residents in emergency setting per day, accounting for about 700 cases annually. As such, this retrospective study aims to focus on evaluating interpretation discrepancies of residents who were at least in their second year of training, as well as reveal the types of commonly discordant lesions that residents encounter difficulty with. A comparison of the different levels of residency was conducted during the study period starting from March 2011 until February 2013, since residents are promoted to the next grade on March 1st every year.

Minor discrepancies were relatively common (25.1%), involving mostly fatigue due to overnight on-call duty or deemed not critical enough for inclusion on the preliminary report based on patients' clinical features (10) (Table 2).

Conversely, intermediate discrepancies involving inadequate description about HIVD including failure to detect extraforaminal disc herniation or pointing out the wrong compressed nerve root level, failure or over-diagnosis of soft tissue injury or fracture of the posterior column may be critical to the patients' clinical course (Table 2).

Because urgent preliminary readings were requested by the emergency room, residents face immense pressure to arrive at an accurate and timely primary decision. Therefore, there is a tendency to over-diagnose especially during trauma evaluation, as evidenced by the most frequent major discrepancy involving fracture over-diagnosis (Table 2, Fig. 1).

Several cases of missed lesion in the main finding or the level of interest (total 81 cases) highlight the careless reading and interpretation of spine MRI by residents. The underlying factors may include fatigue or impetuous attitude for emergency reading. Residents may be poor at the diagnosis of soft tissue injuries such as PLC or DLC leading to missed detection (11 cases), spinal cord abnormality (nine cases), and diffuse marrow abnormality (five cases) (Table 2, Figs. 2, 3). This finding suggests the need for supplementation of the curriculum to meet the cases of spine trauma, spinal cord disease, and bone marrow abnormality many times during the training period.

This study showed a substantially higher discrepancy rate: 25.1% for minor, 6.9% for intermediate, and 7.4% for major types, compared with previous studies (6-8). Siström et al. (6) reported about 2.5% of significant disagreement rate for spine MRI without its clear definition. However, it may be assumed that extremely serious cases related to severe morbidity or mortality were categorized as significant discordant cases, presented at monthly case challenge conferences based on their description. Bruni et al. (7) showed an incidence of 1.2% major discrepancy cases with a significant change inducing negative outcome, but failed to reveal the interpretation error only for spine MRI compared with our results. Filippi et al. (8) reported that 9.2% of minor discrepancies involved spine MRI without any major etiology, significantly lower than our results. These differences in results between this study and previous reports may be attributed to our emphasis on the errors of radiologist report interpretation, rather than its negative clinical impact. However, we tried to demonstrate not only the discrepancy rate of preliminary reports of spine MRI, but also the types of commonly discordant lesions.

Notwithstanding the rigid criteria of interpretation errors, there was no adverse clinical event such as delayed

treatment, similar to the results of a previous study (8), as the preliminary reports were finalized and results conveyed to the referring clinicians within 24 hr.

In our study, there was no significant difference in the incidence of minor, intermediate, and major discrepancies with residents' years of training, especially between second- and third year residents. This finding may be attributed to only the one month-difference in experience involving spine radiology between second- and third year residents, and the few cases read by first- or fourth-year residents. In addition, this study was performed using data spread over three years, therefore, data obtained from the same resident may be accumulated and analyzed together, which may influence the absence of difference according to personal characteristics or cluster effects.

This study has some limitations. First, this study was retrospective and did not analyze the factors affecting interpretation error, such as the working time of the on-call resident or the duration of prior experience in spine radiology. Residents usually focused on the most important MRI finding based on patients' clinical features or questions and may be willing to rest after stating only the main diagnosis. In addition, we only enrolled cases involving the spine MRI performed in the emergency room, which may be insufficient to reveal the weaknesses in the training curriculum of spine radiology compared with analysis of all the cases of spine MRI. However, we assumed that the primary emergency report of spine MRI determined the true conditions representative of residents' training status without support by the staff nearby or the higher grade of residents. Second, the classification of discrepancies into minor, intermediate and major forms is arbitrary and non-standardized, for example a previous study defined only minor and major discrepancies (8). In clinical practice in emergency room, the treatment or management option may be decided promptly without the radiologist report, and it was not proven that whether or not the case with major discrepancy was associated with an adverse outcome in this study because the final report was issued within 24 hr in our hospital. However, this study attempted to demonstrate the common interpretation errors by residents who are relatively inexperienced in reading spine MRIs, and therefore, it was decided to stratify errors based on their probable clinical impact. Third, we did not investigate the type of discrepancy according to the respective indication of spine MRI, due to variation in indications and absence of uniform correlation with the main diagnosis obtained using spine MRI. Fourth, we used a formal report by a single

faculty radiologist, not blinded to the preliminary report as a standard, which may suggest a substantially higher degree of interpretation discrepancy by residents.

In conclusion, there was a 7.4% rate of major discrepancies in preliminary reporting of emergency spinal MRIs made by radiology residents, probably related to a relative lack of clinical experience, indicating the need for further exposure to cases, specific to spine trauma, spinal cord and bone marrow lesions.

REFERENCES

1. Yablon CM, Wu JS, Newman LR, Downie BK, Hochman MG, Eisenberg RL. A needs assessment of musculoskeletal fellowship training: a survey of practicing musculoskeletal radiologists. *AJR Am J Roentgenol* 2013;200:732-740
2. Walls J, Hunter N, Brasher PM, Ho SG. The DePICTORS Study: discrepancies in preliminary interpretation of CT scans between on-call residents and staff. *Emerg Radiol* 2009;16:303-308
3. Chung JH, Strigel RM, Chew AR, Albrecht E, Gunn ML. Overnight resident interpretation of torso CT at a level 1 trauma center an analysis and review of the literature. *Acad Radiol* 2009;16:1155-1160
4. Miyakoshi A, Nguyen QT, Cohen WA, Talner LB, Anzai Y. Accuracy of preliminary interpretation of neurologic CT examinations by on-call radiology residents and assessment of patient outcomes at a level I trauma center. *J Am Coll Radiol* 2009;6:864-870
5. Hochberg AR, Rojas R, Thomas AJ, Reddy AS, Bhadelia RA. Accuracy of on-call resident interpretation of CT angiography for intracranial aneurysm in subarachnoid hemorrhage. *AJR Am J Roentgenol* 2011;197:1436-1441
6. Siström C, Deitte L. Factors affecting attending agreement with resident early readings of computed tomography and magnetic resonance imaging of the head, neck, and spine. *Acad Radiol* 2008;15:934-941
7. Bruni SG, Bartlett E, Yu E. Factors involved in discrepant preliminary radiology resident interpretations of neuroradiological imaging studies: a retrospective analysis. *AJR Am J Roentgenol* 2012;198:1367-1374
8. Filippi CG, Schneider B, Burbank HN, Alsofrom GF, Linnell G, Ratkovits B. Discrepancy rates of radiology resident interpretations of on-call neuroradiology MR imaging studies. *Radiology* 2008;249:972-979
9. Strub WM, Leach JL, Ying J, Vagal A. First year radiology residents not taking call: will there be a difference? *Emerg Radiol* 2007;13:231-235
10. Ruutiainen AT, Durand DJ, Scanlon MH, Itri JN. Increased error rates in preliminary reports issued by radiology residents working more than 10 consecutive hours overnight. *Acad Radiol* 2013;20:305-311