AJNR

Clinical Correlation of a New MR Imaging Method for Assessing Lumbar Foraminal Stenosis

H.-J. Park, S.S. Kim, S.-Y. Lee, N.-H. Park, M.-H. Rho, H.-P. Hong, H.-J. Kwag, S.-H. Kook and S.-H. Choi

ORIGINAL
RESEARCH

H.-J. Park
S.S. Kim
S.-Y. Lee
N.-H. Park
M.-H. Rho
H.-P. Hong
H.-J. Kwag
S.-H. Kook
S.-H. Choi

Clinical Correlation of a New MR Imaging Method for Assessing Lumbar Foraminal Stenosis

BACKGROUND AND PURPOSE: Recently, Lee et al reported a new grading system for the lumbar spinal foraminal stenosis. They considered the type of stenosis, the amount of fat obliteration, and the presence of nerve root compression. Our aim was to evaluate whether a new MR imaging grading system correlated with symptoms and neurologic signs and could replace the previous grading system.

MATERIALS AND METHODS: We examined 91 patients (M/F = 49:42; mean age, 50 years) who visited our institution and underwent MR imaging of the L-spine and were evaluated by 2 musculoskeletal radiologists. The presence and grade of lumbar foraminal stenosis at the maximal narrowing point was assessed according to the new grading system suggested by Lee et al (Lee system) and the Wildermuth grading system (Wildermuth system). Results were correlated with clinical manifestations and neurologic physical examination. Statistical analysis was performed by using κ statistics, categoric regression analysis, and nonparametric correlation analysis (Spearman correlation).

RESULTS: Interobserver agreement in the grading of foraminal stenosis between the 2 readers was substantially correlated (κ of Lee system = 0.767, κ of Wildermuth system = 0.734). The Rs for reader 1 and reader 2 between the Lee system and the Wildermuth system were 0.880 and 0.885, between Lee system and PNM were 0.715 and 0.604, and between the Wildermuth system and PNM were 0.800 and 0.680. For patients younger than 50 years of age, the R between the Lee and Wildermuth systems was higher than that for patients 50 years or older, but the Rs between the grading system and PNM were lower in the younger group than in the older group. The Rs of the Wildermuth system with PNM were higher in the older group than in the younger group; the differences between the Rs of the Lee system with PNM and the Wildermuth system with PNM were higher in the older group (0.016 [young] versus 0.130 [old] and 0.008 versus 0.107).

CONCLUSIONS: Interobserver agreement of the Lee system was slightly higher than the Wildermuth system and substantially correlated. Both systems are good for evaluation of lumbar spinal foraminal stenosis, but the Lee system showed slightly better interobserver agreement and good clinical correlation in the younger group of patients.

ABBREVIATIONS ΔR = differences of R; DTR = deep tendon reflex; EMG = electromyogram; L-spine = lumbar spine; PNM = positive neurologic manifestation; R = correlation coefficient; STIR = short τ inversion recovery

SPINE

ORIGINAL RESEARCH

Lumbar spinal foraminal stenosis is an important pathologic entity to recognize in patients with radicular symptoms. The incidence of lumbar foraminal stenosis is about 10% (8%–11%) in the symptomatic group.[1] MR imaging is considered an appropriate tool for studying spine pathologies, especially spinal stenosis. Wildermuth et al[2] introduced a partially quantitative classification system for grading lumbar spinal foraminal stenosis on the basis of MR imaging findings. They focused on the degree of epidural fat obliteration but did not consider direct nerve root compression or deformity. Recently, Lee et al[3] reported a new grading system for lumbar spinal foraminal stenosis. They considered the type of stenosis, the amount of fat obliteration, and the presence of nerve root compression. The purpose of this study was to evaluate

whether the new MR imaging grading system correlated with symptoms and neurologic signs and could replace the previous grading system.

Materials and Methods

Case Selection
Our study population included 49 (54%) men and 42 (46%) women who visited our institution and underwent MR imaging of the L-spine between January 2010 and September 2010. The age distributions were the following: Forty-nine patients were younger than 50 years (54%), and 42 patients were (46%) older. The mean age was 50.6 years. Exclusion criteria were the following: infection, tumor, acute trauma, surgical history, central disk herniation, combined brain infarction, or other intracranial lesions and peripheral neuropathy. Patients with symptoms at different cord levels or on opposite sides were excluded. This retrospective study was approved by the institutional ethics review board, with no requirement for informed patient consent.

Image Analysis
MR imaging examinations were interpreted by 2 fellowship-trained academic musculoskeletal radiologists who had 12 and 10 years of experience. They were blinded to the clinical information and radio-

logic reports. A total 728 foramina and corresponding nerve roots in 91 patients were qualitatively analyzed from L2–3 to L5–S1. The radiologists assessed the presence and grade of lumbar foraminal stenosis at the maximal narrowing point according to the new grading system suggested by Lee et al.[3] We named this grading system the "Lee system." Lumbar foraminal stenosis was classified in the Lee system on the basis of the T1-weighted sagittal images as follows: grade 0 = absence of foraminal stenosis, grade 1 = mild foraminal stenosis showing perineural fat obliteration surrounding the nerve root in 2 opposing directions without morphologic change of the nerve, grade 2 = moderate foraminal stenosis showing perineural fat obliteration surrounding the nerve root in 4 directions without morphologic change of the nerve, and grade 3 = severe foraminal stenosis showing nerve root collapse or morphologic change. When the radiologists were not confident about their findings, the T2-weighted sagittal and axial images were used as a supplementary evaluation method.

We could exclude the possibility of partial volume artifacts of the lesion from the nerve root itself and could trace the full course of the nerve root through the T2-weighted images. The T2-weighted axial image was especially useful in the evaluation of perineural cysts or nerve root swellings that can cause false-positive results. The radiologists also assessed the foraminal stenosis with the Wildermuth grading system (Wildermuth system) on the basis of the T1-weighted sagittal images. The severity of foraminal stenosis was graded 0 if the foramina were without pathology (normal dorsolateral border of the intervertebral disk and normal form at the foraminal epidural fat); grade 1 if slight foraminal stenosis and deformity of the epidural fat were seen, with the remaining fat still completely surrounding the exiting nerve root; grade 2 if marked foraminal stenosis with epidural fat only partially surrounding the nerve root was seen; and grade 3 for advanced stenosis with obliteration of the epidural fat.

Clinical Correlation

Neurologic examinations were performed and clinical manifestations were acquired by the same physician. We considered neurologic clinical manifestations as observed paresthesia, extremity weakness, numbness, and funicular or radicular pain. Positive neurologic signs were a positive Lhermitte sign, Spurling sign, decreased response of DTR, and a positive denervation sign on EMG. DTR evaluations were performed in all patients, and EMG was performed in 5. More than 1 positive neurologic sign combined with >1 neurologic clinical manifestation was considered a PNM of the lumbar foraminal stenosis. Symptoms of upper extremities or the opposite side of the stenosis were not considered PNMs.

MR Imaging Parameters

All MR imaging examinations were performed by using the same protocol with a 1.5T magnet (Intera; Philips Healthcare, Best, the Netherlands) by using a Syn-Spine coil (Philips Healthcare) and fast spin-echo imaging. T1-and T2-weighted images were obtained in the axial plane, and T1- and T2-weighted images and T2 STIR images were obtained in the sagittal plane in supine position. An FOV of 32 cm, matrix 512 × 256, and section thickness of 4 mm were used for sagittal views; an FOV of 15 cm, matrix 256 × 320, and section thickness of 4 mm were used for axial images. MR images were sagittal T1-weighted spin-echo (TR/TE = 500–600/ 12–17 ms), T2 STIR (TR/TE = 2500/60 ms), T2-weighted (TR/TE, 3500/120 ms), axial T2-weighted turbo spin-echo (TR/TE, 3000–

Table 1: Incidence of grades for the Lee system

	Grade 0	Grade 1	Grade 2	Grade 3	Total
Reader 1	56	19	8	8	91
Reader 2	65	12	8	6	91

Table 2: Incidence of grades for the Wildermuth system

	Grade 0	Grade 1	Grade 2	Grade 3	Total
Reader 1	50	16	18	7	91
Reader 2	60	8	18	3	91

4000/60 ms), and T1-weighted turbo spin-echo (TR/TE, 600–700/ 10–15 ms).

Statistical Analysis

Interobserver agreement between the 2 radiologists was analyzed by using κ statistics. κ value interpretation was the following: poor ($\kappa < 0.1$), slight ($0.1 \leq \kappa \leq 0.2$), fair ($0.2 < \kappa \leq 0.4$), moderate ($0.4 < \kappa \leq 0.6$), substantial ($0.6 < \kappa \leq 0.8$), and almost perfect ($0.8 < \kappa \leq 1$). Rs were calculated with categoric regression analysis and nonparametric correlation analysis (Spearman correlation). For analysis of the relationship between findings and patient characteristics, the association of the MR imaging findings and clinical manifestations was evaluated by age (younger than 50 years, equal to or older than 50 years). R between 0.7 and 0.9 was considered a relatively high correlation, and $R > 0.9$ was a very high correlation. The level of correlation significance was .01. Statistical analyses were performed by using the Statistical Package for the Social Sciences software, Version 10.1 (SPSS, Chicago, Illinois).

Results

With the Lee system, reader 1 detected grade 0 foraminal stenosis in 56 patients, grade 1 in 19 patients, grade 2 in 8 patients, and grade 3 in 8 patients. Reader 2 detected grade 0 in 65 patients, grade 1 in 12 patients, grade 2 in 8 patients, and grade 3 in 6 patients (Table 1). With the Wildermuth system, reader 1 detected grade 0 foraminal stenosis in 50 patients, grade 1 in 16 patients, grade 2 in 18 patients, and grade 3 in 7 patients. Reader 2 detected grade 0 in 60 patients, grade 1 in 8 patients, grade 2 in 18 patients, and grade 3 in 5 patients (Table 2 and Figs 1 and 2). We found PNMs in 42 patients, with negative findings for 49 patients. Mismatched cases with the Lee system and Wildermuth system were 14 patients for reader 1 and 14 patients for reader 2 (Figs 3 and 4). For correlation with the Lee system grade 0, 9 patients (reader 1) and 17 patients (reader 2) were neurologic manifestation−positive; with the Lee system grade 2 or 3; no patients were neurologic manifestation−negative. For the Lee system grade 1, only 2 patients (reader 1) and 1 patient (reader 2) were neurologic manifestation−negative. The following were mismatches between the Lee system and the Wildermuth system: 1 grade difference for 16 patients for reader 1 and 13 patients for reader 2; and >2 grade differences for 1 patient for reader 1 and 1 patient for reader 2.

Interobserver agreement in the grading of foraminal stenosis between the 2 readers was found to be substantially correlated (κ-value for the Lee system = 0.767, κ-value for the Wildermuth system = 0.734). Interobserver agreement for the
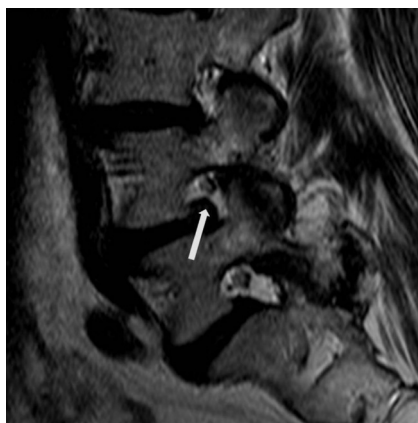
Fig 1. *A* and *B*, The Lee system grade 2 and the Wildermuth system grade 2. T2-weighted images of a 39-year-old man and a 64-year-old man show narrowing of the vertical and transverse width of neural foramina in the left L5–S1 and the right L5–S1. Decreased intervertebral disk space, thickened ligamentum flavum, and disk protrusions are seen (*arrows*). Perineural fat obliteration is also seen, but nerve root deformity is not noted.

Fig 2. The Lee system grade 3 and the Wildermuth system grade 3. T1-weighted image of an 82-year-old woman revealing marked narrowing of the vertical and transverse width of neural foramina at L5–S1. Decreased intervertebral disk space, thickened ligamentum flavum, and disk protrusions are seen (*arrows*). The nerve root is collapsed and deformed compared with another nerve root at a different level.
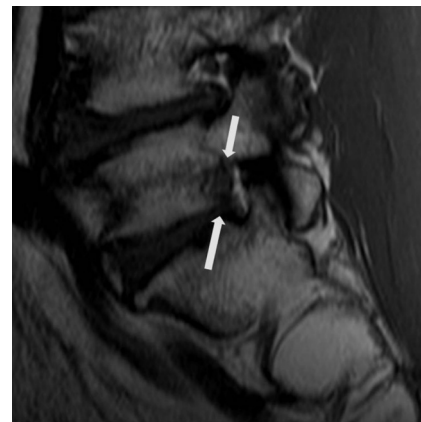
Fig 4. The Lee system grade 1 and the Wildermuth system grade 2. T1-weighted image of a 54-year-old man shows narrowing of the vertical width of the neural foramen and decreased intervertebral disk space in the left L5–S1 (*arrows*). Partial perineural fat obliteration is noted, but deformity of the nerve root is not seen.

Fig 3. The Lee system grade 0 and the Wildermuth system grade 1 foraminal stenosis. T2-weighted sagittal image of a 78-year-old woman shows a normal nerve root without compression but mild narrowing of the foramen in the right L4–5. The arrow indicates the protruded disk.

Table 3: *R*s for lumbar foraminal stenosis[a]

Observer	L System vs W System	L System vs W System	W System vs PNM
Reader 1	0.888	0.751	0.800
Reader 2	0.885	0.751	0.680

Note:—L system indicates Lee system; W system, Wildermuth system.
[a] The level of correlation significance was .01.

new grading system (Lee system) was similar to but slightly higher than that for the Wildermuth system.

R for reader 1 between the Lee system and the Wildermuth system was 0.880, R for the Lee system and PNM was 0.715, and R for the Wildermuth system and PNM was 0.800. R for reader 2 for Lee system and Wildermuth system was 0.885, R for the Lee system and PNM was 0.604, and R for the Wildermuth system and PNM was 0.680 (Table 3). The ∆Rs between the Lee system and PNM and between the Wildermuth system and PNM were 0.085 for reader 1 and 0.076 for reader 2. In the younger age group (younger then 50 years), R for reader 1 between the Lee system and the Wildermuth system was 0.973, R for the Lee system and PNM was 0.711, and R for the Wildermuth system and PNM was 0.695. In the older group (50 years or older), R for reader 1 between the Lee system and Wildermuth system was 0.832, R for the Lee system and PNM was 0.612, and R for the Wildermuth system and PNM was 0.742. In the younger group (younger than 50 years), R for reader 2 between the Lee system and Wildermuth system was 0.985, R for the Lee system and PNM was 0.511, and R for the

Table 4: _R_s for lumbar foraminal stenosis for patients younger than 50 years of age[a]

Observer	L System vs W System	L System vs PNM	W System vs PNM
Reader 1	0.973	0.711	0.695
Reader 2	0.985	0.511	0.503

Note:—L system indicates Lee system; W system, Wildermuth system.
[a] The level of correlation significance was .01.

Table 5: _R_s for lumbar foraminal stenosis for patients older than 50 years of age[a]

Observer	L System vs W System	L System vs PNM	W System vs PNM
Reader 1	0.832	0.612	0.742
Reader 2	0.841	0.530	0.637

Note:—L system indicates Lee system; W system, Wildermuth system.
[a] The level of correlation significance was .01.

Wildermuth system and PNM was 0.503. In the older group (50 years or older), the _R_ for reader 2 for the Lee system and the Wildermuth system was 0.841, _R_ for the Lee system and PNM was 0.530, and _R_ for the Wildermuth system and PNM was 0.637 (Tables 4 and 5). The Δ_R_s between the Lee system and PNM and the Wildermuth system and PNM were 0.107–0.130 in the older group and 0.008–0.016 in the younger group.

The Lee system was highly correlated with the Wildermuth system, especially in the older group of patients. The correlation with neurologic manifestation was higher in the Wildermuth system, but in the younger group, the Lee system showed slightly higher correlation with neurologic manifestations.

Discussion

The lateral nerve root canal is a tubular-shaped region in which the nerve root passes from the thecal sac to the intervertebral foramen. Jenis et al[4] defined the lumbar intervertebral foramen as a vertical interpedicular zone incorporating portions of the lateral recess and exit zone. Thus, the boundaries include the superior and inferior pedicles, the posteroinferior margin of the superior vertebral body, the posterolateral margin of the intervertebral disk, and the posterosuperior margin of the inferior vertebral body anteriorly. The posterior boundaries are the ligamentum flavum and the superior articular facet. Foraminal stenosis might result in significant radicular pain from compression of structures within the area.

Previous reports have suggested the symptom mechanism of lumbar foraminal stenosis. Hasegawa et al[5] investigated critical heights of the intervertebral disks and foramina and concluded that critical dimensions might be indicators of foraminal stenosis. As the superior facet continues to subluxate, the alteration of biomechanical forces contributes to the development of a hypertrophic ligamentum flavum and bony spurs, which may diminish the volume of the foramen to a greater extent. The combination of disk space narrowing and overgrowth of the facet joint capsule may lead to transverse foraminal stenosis (transverse stenosis). The exiting nerve root is compressed between the superior articular facet and the posterior vertebral body in a transverse direction. Another cause of foraminal stenosis is craniocaudal compression (vertical stenosis). Posterolateral osteophytes from the vertebral endplates protrude into the foramen along with a laterally bulging annulus or herniated disk, compressing the nerve root against the superior pedicle. These static changes in foraminal volume may develop and cause circumferential stenosis.[6]

Wildermuth et al[2] introduced a semiquantitative classification in 1998, focusing on the degree of epidural fat obliteration and the absence of characteristics of morphologic nerve root change. Attias et al[1] summarized the Wildermuth classification system as having poor interobserver reliability with significant differences between foraminal measurements performed by MR imaging and on cadavers. Recently, Lee et al[3] suggested a new MR imaging grading system including both perineural fat obliteration and nerve root morphology based on sagittal MR images. They reported that the system is more practical for grading foraminal stenosis, given the frequency of radiculopathy caused by nerve root irritation.

In our study, the interobserver agreement for the Wildermuth system was slightly lower than that for Lee system, but the difference was minimal. The interobserver agreement for the Lee system was slightly lower than that of the study by Lee et al[3] (0.767 versus 0.909–0.942). We found interobserver agreement to be substantially, but not perfectly, correlated. We also correlated the new system with clinical manifestation and neurologic signs. _R_ between the Lee system and the Wildermuth system was high (0.880, 0.885) and similar. _R_ for the grading systems and PNMs were different, with the Wildermuth system more closely correlated than the Lee system (0.800 and 0.680 versus 0.715 and 0.604). We grouped study subjects by age. In the younger group (younger than 50 years), _R_ between the Lee system and the Wildermuth system was higher than that in the older group (50 years or older) (0.973 and 0.985 versus 0.832 and 0.841), but _R_s for the grading system and PNM were lower for the older group (0.711 and 0.511 versus 0.612 and 0.530 in the Lee system and PNM, 0.695 and 0.503 versus 0.742 and 0.637 in the Wildermuth system and PNM). _R_s for Wildermuth system with PNM were higher in the older group and lower in the younger group, and differences between _R_s for the Lee system with PNM and the Wildermuth system with PNM were higher in the older group (0.016 and 0.008 versus 0.130 and 0.107). These results suggested that the Wildermuth system reflected clinical symptoms more precisely, especially for the older age group (Δ_R_ = 0.107, 0.130). In the younger group, the Lee system more closely correlated, but the difference was minimal (Δ_R_ = 0.008, 0.016).

One limitation of this study was the single posture of the lumbar spine MR imaging because lumbar spine posture affects the dimensions of the intervertebral foramina. Mayoux-Benhamou et al[7] found that all diameters of the lumbar foramina decreased as the spine moved into full extension; as the spine moved from flexion to extension, the heights of the foramina were reduced by approximately 18%. In our study, the spines were in a neutral supine position, and no flexion or extension was applied. However, all patients were in the same position, so differences should have been minimal. Another limitation was that the study was not based on a quantitative evaluation of the clinical manifestations. We classified clinical results as either positive or negative, so the dependent variables were unordered qualitative variables. Nonetheless, the purpose of this study was to evaluate the clinical effectiveness of the new grading system, because differentiation of the fo-

raminal stenosis in symptomatic and asymptomatic conditions might be valuable.

Conclusions

Interobserver agreement for the new grading system (the Lee system) was similar to but slightly higher than that for the Wildermuth system and substantially correlated. The Lee system is highly correlated with the Wildermuth system, especially in the older group of patients. The correlation with neurologic manifestations is higher in the Wildermuth system; but in the younger group, the Lee system showed slightly higher correlation with neurologic manifestation. Both systems are good for the evaluation of lumbar spinal foraminal stenosis, but the Lee system showed slightly better interobserver agreement and good clinical correlation in the younger group of patients.

References

1. Attias N, Hayman A, Hipp JA, et al. **Assessment of MRI in the diagnosis of lumbar spine foraminal stenosis: a surgeon's perspective.** *J Spinal Discord Tech* 2006;19:249–56
2. Wildermuth S, Zanetti M, Duewell S, et al. **Lumbar spine: quantitive and qualitative assessment of positional (upright flexion and extension) MR imaging and myelography.** *Radiology* 1998;207:391–98
3. Lee S, Lee JW, Yeon JS, et al. **A practical MRI grading system for lumbar foraminal stenosis.** *AJR Am J Roentgenol* 2010;194:1095–98
4. Jenis LG, An HS, Gordin R. **Foraminal stenosis of the lumbar spine: a review of 65 surgical cases.** *Am J Orthop (Belle Mead NJ)* 2001;30:205–11
5. Hasegawa T, An HS, Haughton VM, et al. **Lumbar foraminal stenosis: critical heights of the intervertebral discs and foramina: a cryomicrotome study in cadaver.** *J Bone Joint Surg Am* 1995;77:32–38
6. Jenis LG, An HS. **Spine update: lumbar foraminal stenosis.** *Spine* 2000;25: 389–94
7. Mayoux-Benhamou MA, Revel M, Aaron C, et al. **A morphological study of the lumbar foramen: influence of flexion-extension movements and of isolated disc collapse.** *Surg Radiol Anat* 1989;11:97–102