The British Journal of Radiology, *79* (2006), 372–377

Observer variation in the evaluation of lumbar herniated discs and root compression: spiral CT compared with MRI

[1]J C VAN RIJN, MD, [2]N KLEMETSO, MD, [1]J B REITSMA, MD, PhD, [1]P M BOSSUYT, PhD, [2]F J HULSMANS, MD, PhD, [4]W C PEUL, MD, PhD, [2]G J DEN HEETEN, MD, PhD, [3]J STAM, MD, PhD and [2]C B L M MAJOIE, MD, PhD

Departments of [1]Clinical Epidemiology and Biostatistics, [2]Radiology and [3]Neurology, Academic Medical Center, Meibergdreef 9, 1105 AZ, Amsterdam and the [4]Department of Neurosurgery, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, The Netherlands

ABSTRACT. Spiral CT is considered the best alternative for MRI in the evaluation of herniated discs. The purpose of this study was to compare radiological evaluation of spiral CT with MRI in patients suspected of herniated discs. 57 patients with lumbosacral radicular syndrome underwent spiral CT and 1.5 T MRI. Two neuroradiologists independently evaluated 171 intervertebral discs for herniation or ''bulge'' and 456 nerve roots for root compression, once after CT and once after MRI. We compared interobserver agreement using the kappa statistic and we performed a paired comparison between CT and MRI. For detection of herniated or bulging discs, we observed no significant difference in interobserver agreement (CT kappa 0.66 *vs* MRI kappa 0.71; $p=0.40$). For root compression, we observed significantly better interobserver agreement at MRI evaluation (CT kappa 0.59 *vs* MRI kappa 0.78; $p=0.01$). In 30 of 171 lumbar discs (18%) and in 54 of 456 nerve roots (12%), the observers disagreed on whether CT results were similar to MRI. In the cases without disagreement, CT differed from MRI in 6 discs (3.5%) and in 3 nerve roots (0.7%). For radiological evaluation of lumbar herniated discs, we found no evidence that spiral CT is inferior to MRI. For evaluating lumbar nerve root compression, spiral CT is less reliable than MRI.

Received 16 August 2005
Revised 23 November 2005
Accepted 9 December 2005

DOI: 10.1259/bjr/26216335

In a series of comparative radiological studies on the evaluation of lumbar herniated discs, MRI was found to be ''slightly better'' than [1–4] or equal to [5–8] CT. We found one study suggesting CT to be more accurate than MRI [9], but overall, MRI is preferred and CT seems to have comparable capacity to diagnose herniated discs. Therefore, current guidelines designate MRI as the first-choice investigation and suggest CT as the alternative in the evaluation of the lumbar back if MRI is contra-indicated or unavailable [11–13].

Besides herniated discs, the direct evaluation of nerve roots by MRI has been considered an important asset to facilitate decision making in patients with back pain [14–17]. No evidence, however, is available on the capabilities of spiral CT to evaluate lumbar nerve root compression.

The purpose of our study was to compare spiral CT with MRI in a series of patients suspected of lumbar herniated discs. For each technique we assessed inter-observer agreement in detecting herniated discs, bulging discs and nerve root compression. We also performed a paired, direct comparison between spiral CT and MRI.

Methods

This study was conducted between June 1999 and June 2000, as part of a larger project on the diagnostic process of patients with lumbosacral radicular pain at the University of Amsterdam, The Netherlands. The Institutional Review Board approved the study protocol.

Patients were recruited from the neurology outpatient department. Those eligible were patients referred by their general practitioner with lumbosacral radicular syndrome (LRS) with suspected disc herniation at the levels L3–L4 to L5–S1 in whom conservative treatment had been unsuccessful. LRS was defined according to the national general practitioners' guideline and the consensus statement on diagnosis and treatment of LRS defined by the Dutch Neurology Society [18]. The hallmark of this definition is continuous monoradicular or multiradicular pain below the knee with a primary suspicion of disc herniation. Excluded were patients younger than 18 years or older than 70 years, pregnant women, patients with a previous history of lumbosacral herniation or lumbosacral surgery, as well as patients with contraindication for MRI. Eligible patients received written and oral information about the study.

After the neurologist had confirmed the diagnosis of LRS, consenting patients were subjected to spiral CT and MRI within 1 week. No specific treatment was given within this period.

Imaging techniques

Spiral CT

Lumbar CT examinations were performed on a 2-slice CT-Twin scanner (Philips Medical Systems, Best, The Netherlands). Helical CT-scans were made with 120 kV, 265 mAs (effective dose: 6.4 mSv), table feed of 1.4 mm s^{-1}, 2 mm × 1 mm collimation (effective slice width, 1.1 mm) and 0.5 mm increment from the level L3 to the bottom of S1. The gantry angle was aligned through the disc space of L4–L5. In addition to the original axial images with the bone window (level 400 HU/window 1600 HU), reformatted 4 mm axial sections of all scanned interspaces were made parallel to the interspaces using the soft tissue window (level 50 HU/window 180 HU). Also reformatted ''0 mm'' (*i.e.* slice width in the order of 1 mm) sagittal sections were made using the bone window, and 4 mm sagittal sections using the soft tissue window.

The quality of the CT-images used in this study, obtained with a 2-slice CT-scanner, is comparable with the quality of more modern, multislice CT-scanners. The effective slice width in this study was 1.1 mm, which is only slightly more than the effective slice width of multislice CT-scanners. The effective slice width is usually 25–30% higher than the collimation width and thus in the order of 0.9–1.0 mm for, *e.g.* a collimation of 16 mm × 0.75 mm [19, 20].

MRI

Lumbar MR examinations were performed with a 1.5 T Signa LX Scanner (General Electric Medical Systems, Milwaukee, WI) using a dedicated lumbar spine surface coil. The protocol included sagittal spin-echo T_1 weighted (repetition time (TR) 500 ms, echo time (TE) 14 ms) and proton density/T_2 weighted (TR 3500 ms, TE 20–120 ms) fast spin echo images with 4 mm slice thickness, 0.5 mm intersection gap, 200 × 512 matrix and 29 cm × 29 cm field of view. In addition, axial spin-echo T_1 weighted (TR 520 ms, TE 12 ms) and fast spin-echo T_2 weighted (TR 4500 ms, TE 120 ms) images were obtained from the increment of L3 to the bottom of S1 with 4 mm slice thickness, 0.5 mm intersection gap, 200 × 256 matrix and 15 cm × 15 cm field of view. Axial images were obtained without angulation.

Image evaluation

Two experienced neuroradiologists (CB and FJ) independently evaluated all CT and MR images with knowledge of the side and level of symptoms. Per observer, the images were evaluated in two independent sessions. There was at least 6 weeks between these sessions. The images of either CT or MRI were presented per patient. It was randomly allocated whether CT or MRI was evaluated in the first session.

Three lumbar discs were examined per patient at levels L3–L4 to L5–S1. Each disc was scored for the presence of a herniation. No distinction between protrusion and extrusion was made; both were considered ''herniated disc''. If no herniation was detected, readers evaluated the presence of a bulging disc. The definition of a bulging disc was according to the description by Jensen [21]: ''circumferential symmetric extension of the disc beyond the interspace''.

Nerve roots L3 through S1 were evaluated per side (eight nerve roots per patient) within the scanned area. A nerve root was defined as the part of the nerve between the central canal and the ganglion. As the scan protocol for both techniques was from the increment of vertebra L3 through the bottom of S1, the radiologists were not able to evaluate a minimal proximal part of root L3 and a minimal distal part of root S1.

A five-point scale was used per nerve root, anchored as definitely no root compression, possibly no root compression, indeterminate, possibly root compression, and definitely compression. For the analyses, these responses were dichotomized as ''root compression'' (possibly or definitely) or ''no root compression'' (all other categories). MRI examples of the different categories have been presented in an earlier report on the same series of patients [22].

Data analysis

Per technique, we calculated interobserver agreement for the evaluation of herniated discs and bulging discs by constructing 3 × 3 tables. Data were matched for lumbar level (L3–L4 to L5–S1). Presence of nerve root compression was analysed using a 2 × 2 table. These data were matched for each side (left and right) and for each nerve root (L3 to S1).

To acknowledge possible correlation between intervertebral discs and nerve roots within one patient, bootstrap sampling was used to calculate the unweighted interobserver kappa statistics and standard-errors for each contingency table [23]. A two-sided *z*-test was then used to compare the kappa statistics of CT with MRI.

To differentiate between CT versus MRI differences due to disagreement and ''true'' differences (no disagreement amongst observers), we performed a paired comparison of the results of CT and MRI. We limited the analysis to the detection of herniated discs. Therefore, the readings for each technique were recoded to ''herniated disc'' or ''no herniated disc''. The latter category comprised bulging discs and discs reported as having no abnormality.

The results were coded per technique as follows: (1) both observers agreed on the presence of a Hernia Nuclei Pulposi (HNP), (2) observers disagreed or (3) both observers agreed on the absence of a HNP. Subsequently, a 3 × 3 table including the results of both observers was constructed to compare both techniques on a per patient basis. The McNemar test for paired data was used to detect differences between modalities.

All calculations were performed with SAS 9.0 and SPSS 11.0. Values lower than $p > 0.05$ were considered to indicate statistical significance.

Results

64 consecutive eligible patients with LRS were identified. Three patients did not undergo MRI because of claustrophobia. Four more patients had incomplete data of spiral CT and MRI, leaving 57 patients with 171 lumbar discs and 456 lumbar roots for analysis.

Table 1 shows the interobserver agreement. There was no significant difference between CT and MRI with respect to interobserver agreement of herniated disc and bulging disc evaluation, but for root compression the kappa statistic was significantly higher for MRI. Figure 1 shows an example of a case in which CT and MRI agree on a HNP with root compression without any interobserver disagreement.

Table 2 shows the results of the paired comparison of CT versus MRI in detecting herniated discs and root compression. Of 171 discs, observers disagreed on herniation at CT evaluation in 20 discs (12%) compared with 12 discs (7%) at MRI. The difference was not significant (McNemar, p=0.22).

In 141 (83%) cases there was no disagreement at CT evaluation as well as no disagreement at MRI evaluation. In three of these cases both readers reported a HNP at CT, but not at MRI evaluation, and three times both readers reported a HNP at MRI but not at CT evaluation.

Observers disagreed on the presence of root compression when evaluating CT in 40 of 456 nerve roots (8.8%) compared with 22 roots (4.8%) when evaluating MRI. The difference was significant (McNemar p=0.03). An example of a case in which observers disagreed on root compression at CT but not on MRI is presented in Figure 2.

In 402 roots (88%) there was no disagreement at CT evaluation as well as no disagreement at MRI evaluation. Of these cases, twice both readers reported root compression at CT but not at MRI evaluation. In one root, both readers reported root compression at MRI but not at CT evaluation (case is presented in Figure 3).

Discussion

In this study we found no evidence that spiral CT is inferior to MRI in the evaluation of herniated lumbar discs and bulging discs, but we observed better interobserver agreement at MRI with respect to the evaluation of nerve root compression.

The paired comparison revealed that in the vast majority of the observed differences between spiral CT and MRI there was interobserver disagreement involved (either at CT or at MRI or at both evaluations). In these cases, we cannot discriminate between differences due to reader variability or due to ''real'' difference between modalities.

The observed differences between CT and MRI without any disagreement indicate possible ''real'' differences. However, after re-evaluation of these cases we can conclude that all HNPs and all cases of root compression were detectable on the CT images as well as on the MR images.

It remains impossible to determine which one of these imaging techniques corresponded best with ''the truth'' because there is no proper reference standard available for this study. However, we observed no particular pattern in the distribution of discordant readings: the differences between spiral CT and MRI were divided equally amongst both techniques. This supports the assumption that spiral CT and MRI are comparable in herniated disc evaluation.

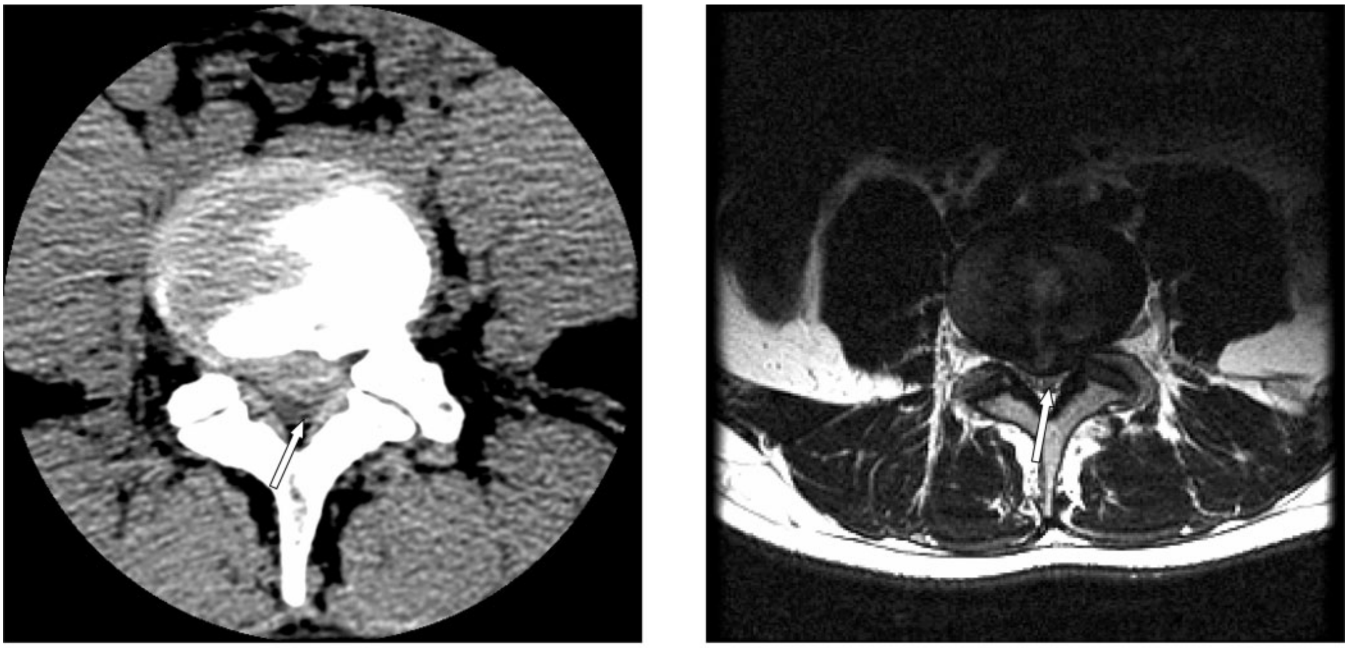
Observer disagreement is a general obstacle in the radiological evaluation of lumbar intervertebral discs, both for spiral CT as well as for MRI. In an earlier published study in the same series of patients, we

Table 1. Spiral CT and MRI interobserver kappa-statistics with 95% confidence intervals

Lumbar disc evaluation									
Spiral CT					**MRI**				
	Observer 2					**Observer 2**			
Observer 1	**HNP**	**Bulge**	**No HNP**	**Total**	**Observer 1**	**HNP**	**Bulge**	**No HNP**	**Total**
HNP	49	10	6	65	HNP	52	3	1	56
Bulge	3	24	9	36	Bulge	7	20	4	31
No HNP	1	9	60	70	No HNP	1	15	68	84
Total	53	43	75	171	Total	60	38	73	171
Kappa 0.66 (0.56–0.75)					Kappa 0.71 (0.63–0.80)				
Difference=0.05; p=0.40									

Root compression							
Spiral CT				**MRI**			
	Observer 2				**Observer 2**		
Observer 1	**RC**	**No RC**	**Total**	**Observer 1**	**RC**	**No RC**	**Total**
RC	35	16	51	RC	45	10	55
No RC	24	381	405	No RC	12	389	401
Total	59	397	456	Total	57	399	456
Kappa 0.59 (0.47–0.70)				Kappa 0.78 (0.69–0.87)			
Difference=0.19; p=0.01							

HNP, Hernia Nuclei Pulposi; RC, root compression.



(a) (b)

Figure 1. 22-year-old male with left-sided lumbosacral radicular syndrome of L5 and paresis of the anterior tibial muscle and with hypaesthesia of L5 and S1. Images show a left-sided herniated disc at level L4–L5 with root compression of L5 reported at CT as well as at MRI by both observers (arrows). (a) Axial CT L4–L5. (b) Axial T_2 weighted MRI L4–L5.

investigated the possible causes of interobserver disagreement in MRI evaluation [22]. We identified a lack of consensus on nomenclature of bulging discs as the main cause of disagreement. Other causes were additional pathology (*e.g.* spondylolisthesis or collapsed vertebral bodies) and small sized lesions, which were believed to be missed more easily.

The disagreement on nerve root evaluation can for a large part be attributed to the lack of a proper standardized and reproducible method to categorize root compression. We chose to use a five point scale to evaluate root compression. Compared with a dichotomous test result, the advantage of this method was that it resembled clinical practice because the observers were given the opportunity to express their uncertainty. A limitation of this method was that we were not able to retrospectively retrieve the decisive items to get to a radiological diagnosis.

Future studies should focus on reducing observer disagreement through the development of standardized radiological nomenclature for lumbar back abnormalities. Development of universal selection criteria to identify specific target groups of patients with back pain and more specific imaging strategies are needed to improve efficiency in decision making. This is also supported by the recent MR-based study by

Table 2. Paired comparison of spiral CT versus MRI

Herniated disc detection				
	MRI			
CT	**HNP**	**Disagree**	**No HNP**	**Total**
HNP	43	4	3	50
Disagree	6	2	12	20
No HNP	3	6	92	101
Total	52	12	107	171
Root compression				
	MRI			
CT	**RC**	**Disagree**	**No RC**	**Total**
RC	32	1	2	35
Disagree	12	8	20	40*
No RC	1	13	367	381
Total	45	22*	389	456

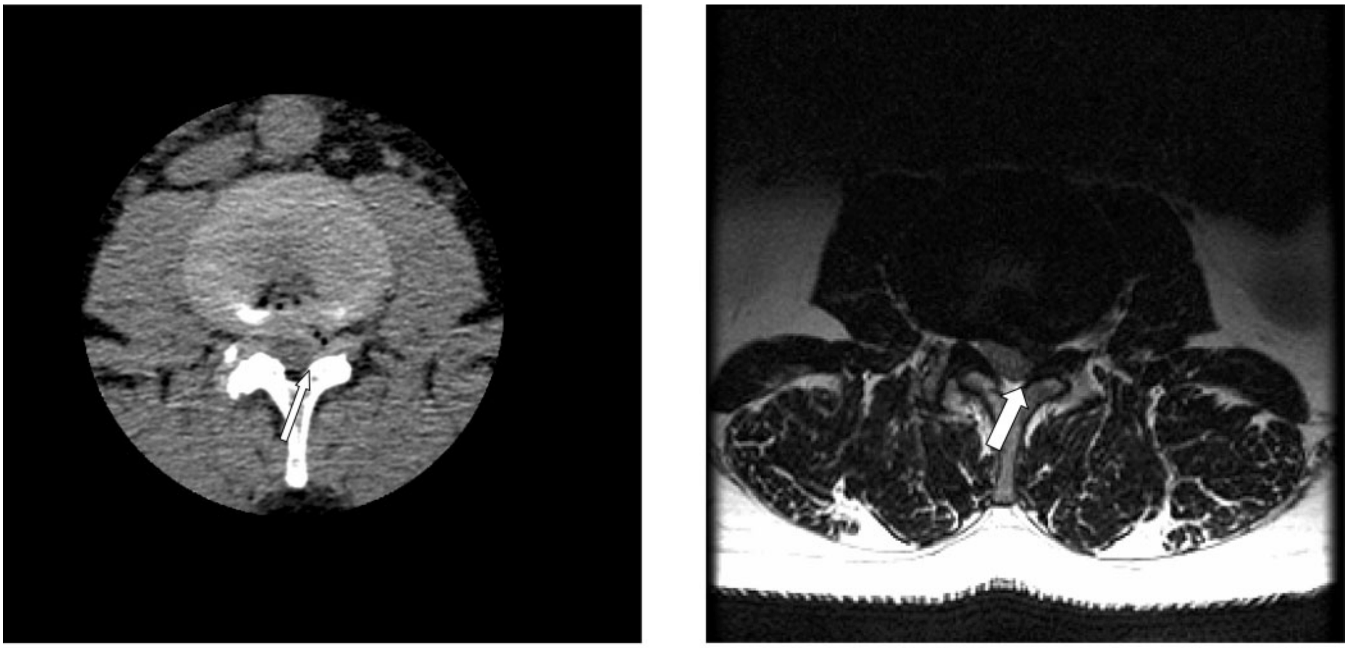
HNP, both observers agreed on the presence of a HNP; No HNP, both observers agreed on the absence of a HNP.
*Significantly more disagreement at CT evaluation (McNemar $p=0.03$).
HNP, Hernia Nuclei Pulposi; RC, root compression.

(a) (b)

Figure 2. 43-year-old male with left-sided lumbosacral radicular syndrome of S1. At both CT and MRI there is disagreement on root compression of L5 on the opposite side of signs and symptoms. (a) Axial CT L4–L5. Observers disagree on whether root L5 on the right (asymptomatic side), is compressed (arrow). (b) Axial T_2 weighted MRI L4–L5. Observers agree on compression of L5 on the right (asymptomatic side) (arrow).

Cihangiroglu in 2004 [24]. To investigate the additional value of myelography (either conventional or by CT or MR imaging) might also be a valuable goal for future studies. Especially in the evaluation of nerve root impingement, it is suggested that myelography can play an important role [25].

In our study, the observers were not blinded for side and level of symptoms. This was done to mimic practice. As part of a larger project, we also obtained blinded evaluations of the same images. We observed no significant differences between blinded and not blinded evaluations [22]. After re-evaluation of the discordant



(a) (b)

Figure 3. 58-year-old male with left-sided lumbosacral radicular syndrome of L3. (a) Axial CT L3–L4. Both observers reported ''definitely no compression'' of L3 on the left (arrow). (b) Axial T_2 weighted MRI L3–L4. Both observers reported ''definitely compression'' of L3 on the left (arrow).

results in the present study, we have no reason to believe that the evaluations were influenced by knowledge of suspected side and level of the LRS.

Our study was limited to the evaluation of herniated discs, bulging discs and root compression. MRI has better qualities to depict all surrounding soft tissue in the lumbar area. This property of MRI is one of the major reasons for neurosurgeons to use MRI to determine whether a patient is a candidate for surgical intervention. We did not investigate to what extent spiral CT might be useful within surgical triage.

In summary, we recommend MRI as the investigation of choice in patients suspected of herniated discs. Nevertheless, a significant number of patients are not able to undergo MRI due to claustrophobia or other contraindications [26, 27]. In these patients, spiral CT seems to be an excellent alternative. However, for the evaluation of root compression, spiral CT appears to be not as reliable as MRI.

References

1. Goscinski I, Ulatowski S, Urbanik A. Comparison of the clinical usefulness of magnetic resonance (MR), computer tomography (CT) and radiculography (R) in diagnostic lumbar discopathy. Przegl Lek 2001;58:885–8.
2. Janssen ME, Bertrand SL, Joe C, et al. Lumbar herniated disk disease: comparison of MRI, myelography, and post-myelographic CT scan with surgical findings. Orthopedics 1994;17:121–7.
3. Forristall RM, Marsh HO, Pay NT. Magnetic resonance imaging and contrast CT of the lumbar spine. Comparison of diagnostic methods and correlation with surgical findings. Spine 1988;13:1049–54.
4. Pevsner PH, Ondra S, Radcliff W, et al. Magnetic resonance imaging of the lumbar spine. A comparison with computed tomography and myelography. Acta Radiol Suppl 1986;369:706–7.
5. Albeck MJ, Hilden J, Kjear L, et al. A controlled comparison of myelography, computed tomography and magnetic resonance imaging in clinically suspected lumbar disc herniation. Spine 1995;20:443–8.
6. Tullberg T, Grane P, Rydberg J, et al. Comparison of contrast-enhanced computed tomography and gadolinium-enhanced magnetic resonance imaging one year after lumbar discectomy. Spine 1994;19:183–8.
7. Thornbury JR, Fryback DG, Turski, PA, et al. Disk-caused nerve compression in patients with acute low-back pain: diagnosis with MR, CT myelography and plain CT. Radiology 1993;186:731–8.
8. Jackson RP, Cain JE, Jacobs RR, et al. The neuroradiographic diagnosis of lumbar herniated nucleus pulposus: II. A comparison of computed tomogrpahy (CT), myelography, CT-myelography, and magnetic resonance imaging. Spine 1989;14:1362–7.
9. Dellerud R, Johansen JG, Johnsen UL, et al. Differentiation between contained and noncontained lumbar disk hernias by CT and MR imaging. Acta Radiol 1995;36:491–6.
10. Eysel P, Rompe JD, Schaub T, et al. Value of imaging techniques of lumbar intervertebral disk prolapse. Z Orthop Ihre Grenzgeb 1994;132:371–6.
11. Patel N. Surgical disorders of the thoracic and lumbar spine: a guide for neurologists. J Neurol Neurosurg Psychiatry 2002;73(Suppl. I):i42–8.
12. Milette PC. Classification, diagnostic imaging, and imaging characterization of a lumbar herniated disk. Radiol Clin N Am 2000;38:1267–92.
13. Herzog RJ. The radiologic assessment for a lumbar disc herniation. Spine 1996;21(24S):19S–38S.
14. Vroomen PC, De Krom MC, Wilmink JT, et al. Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. J Neurol Neurosurg Psychiatry 2002;72:630–4.
15. Pfirrmann CW, Dora C, Schmid MR, et al. MR Image-based grading of lumbar nerve root compromise due to disk herniation: reliability study with surgical correlation. Radiology 2004;230:583–8.
16. Nygaard OP, Jacobsen EA, Solberg T, et al. Nerve root signs on postoperative lumbar MR imaging. A prospective cohort study with contract enhanced MRI in symptomatic and asymptomatic patients one year after microdiscectomy. Acta Neurochir (Wien) 1999;141:619–23.
17. Jarvik JJ, Hollingworth W, Haegerty P, et al. The longitudinal assessment of imaging and disability of the back (LAIDBack) study. Spine 2001;26:1158–65.
18. Stam J. Consensus on diagnosis and treatment of lumbosacral root entrapment syndromes. Ned Tijdschr Geneeskd 1996;140:52.
19. Prokop M. General principles of MDCT. Eur J Radiol 2003;45:S4–S10.
20. Prokop M. Multislice CT: technical principles and future trends. Eur Radiol 2003;13 (Suppl. 5):M3–13.
21. Jensen MC, Brant-Zawadski MN, Obuchowski N, et al. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;331:69–73.
22. Van Rijn JC, Klemetso N, Reitsma JB, et al. Observer variation in MRI evaluation of patients suspected of lumbar disk herniation. AJR Am J Roentgenol 2005;184:299–303.
23. Efron B, Tibshirani RJ. An introduction to the bootstrap. Chapman & Hall, 1993.
24. Cihangiroglu M, Yildirim H, Bozgeyik Z, et al. Observer variability based on the strength of MR scanners in the assessment of lumbar degenerative disc disease. Eur J Radiol 2004;51:202–8.
25. Bartynski WS, Lin L. Lumbar root compression in the lateral recess: MR imaging, conventional myelography, and CT myelography comparison with surgical confirmation. AJNR Am J Neuroradiol 2003;24:348–60.
26. Melendez JC, McCrank E. Anxiety-related reactions associated with magnetic resonance imaging examinations. JAMA 1993;270:745–7.
27. Kanal E, Shellock FG, Talagala L. Safety considerations in MR imaging. Radiology 1990;176:593–606.