

# Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic systematic review

Merel Wassenaar · Rogier M. van Rijn · Maurits W. van Tulder ·  
Arienne P. Verhagen · Danielle A. W. M. van der Windt · Bart W. Koes ·  
Michiel R. de Boer · Abida Z. Ginai · Raymond W. J. G. Ostelo

Received: 24 February 2011 / Revised: 27 July 2011 / Accepted: 31 August 2011 / Published online: 16 September 2011  
© Springer-Verlag 2011

## Abstract

**Purpose** In about 5% of all cases LBP is associated with serious underlying pathology requiring diagnostic confirmation and directed treatment. Magnetic resonance imaging (MRI) is often used for this diagnostic purpose yet its role remains controversial. Consequently, this review

aimed to summarize the available evidence on the diagnostic accuracy of MRI for identifying lumbar spinal pathology in adult low back pain (LPB) or sciatica patients. **Methods** MEDLINE, EMBASE and CINAHL were searched (until December 2009) for observational studies assessing the diagnostic accuracy of MRI compared to a reference test for the identification of lumbar spinal pathology. Two reviewers independently selected studies for inclusion, extracted data and assessed methodological quality. Pooled summary estimates of sensitivity and specificity with 95% confidence intervals were calculated for homogenous subsets of studies.

**Results** Eight studies were included in this review. Strata were defined for separate pathologies i.e. lumbar disc herniation (HNP) and spinal stenosis. Five studies comparing MRI to findings at the surgery for identifying HNP were included in a meta-analysis. Pooled analysis resulted in a summary estimate of sensitivity of 75% (95% CI 65–83%) and specificity of 77% (95% CI 61–88%). For spinal stenosis pooling was not possible.

**Conclusions** The results suggest that a considerable proportion of patients may be classified incorrectly by MRI for HNP and spinal stenosis. However, the evidence for the diagnostic accuracy of MRI found by this review is not conclusive, since the results could be distorted due to the limited number of studies and large heterogeneity.

**Keywords** Diagnostic accuracy · Systematic review · MRI · Low back pain · Diagnostic imaging

## Abbreviations

LBP Low back pain  
MRI Magnetic resonance imaging  
HNP Herniated nucleus pulposus

---

M. Wassenaar · M. W. van Tulder · M. R. de Boer ·  
R. W. J. G. Ostelo  
Department of Health Sciences, Faculty of Earth and Life  
Sciences, VU University Amsterdam, Amsterdam,  
The Netherlands

M. Wassenaar (✉)  
Department of Neurology and Neurosurgery, University Medical  
Center Utrecht, Rudolf Magnus Institute of Neuroscience,  
Room KG01.306.0, Heidelberglaan 100, P.O. box 85500,  
3508, GA, Utrecht, The Netherlands  
e-mail: m.wassenaar@umcutrecht.nl

R. M. van Rijn · A. P. Verhagen · B. W. Koes  
Department of General Practice, University Medical Center,  
Erasmus MC, Rotterdam, The Netherlands

M. W. van Tulder · R. W. J. G. Ostelo  
Department of Epidemiology and Biostatistics, EMGO Institute  
for Health and Care Research, VU University Medical Center  
Amsterdam, Amsterdam, The Netherlands

D. A. W. M. van der Windt  
Arthritis Research UK Primary Care Center,  
Primary Care Sciences, Keele University,  
Staffordshire ST5 5BG, United Kingdom

A. Z. Ginai  
Department of Radiology, University Medical Center,  
Erasmus MC, Rotterdam, The Netherlands

## Introduction

Low back pain, (LBP), defined as pain in the lumbar spinal region with or without sciatica, is a common cause of disability worldwide, with a lifetime prevalence of 60–85% [1–3]. In about 95% of all cases LBP is nonspecific, yet it can be caused by serious underlying pathology such as; disc herniation, spinal stenosis, infection, inflammation, tumour or fractures [4, 5]. In case of suspicion of serious spinal pathology, diagnostic confirmation is required since delayed treatment has been associated with poorer outcomes [1, 4, 5]. Despite the guideline recommendations, diagnostic confirmation is also often requested when the likelihood of a specific cause of LBP is very low [4, 6]. This is often due to the fear of missing serious pathology or to reassure patients [7]. Diagnostic confirmation can be obtained by several imaging techniques one among which is magnetic resonance imaging (MRI). These imaging results combined with available clinical information aid the physicians in their treatment decisions [8, 9]. Of all available techniques MRI is currently the imaging modality of choice. MRI has the advantage of not using ionising radiation and has good visualizing capacities especially of the soft tissues. Thereby it is regarded the most useful method for the detection of spinal infections, spinal metastases, nerve root disorders and disc abnormalities [10]. Nonetheless, the role of MRI in diagnosing lumbar spinal pathology remains controversial [11, 12]; Partly because many studies did not report differences in patient outcomes when comparing lumbar MRI and subsequent treatment to conservative care without diagnostic imaging. However, most studies included patients with a low risk of serious underlying pathology [13–15]. Furthermore, studies have not provided conclusive information about the diagnostic accuracy of MRI [16]. This is largely explained by the absence of a ‘gold’ reference standard for identifying serious underlying spinal pathology in LBP [13, 17, 18]. Additionally heterogeneity in primary diagnostic studies complicates the interpretation of results of diagnostic test accuracy. Potential sources of heterogeneity include variation in; considered pathologies, MRI techniques, reference standards, patient population and methodological quality. To provide more evidence on the diagnostic role of MRI in LBP and sciatica, this systematic review aims to summarize the available evidence on its diagnostic accuracy in the identification of serious underlying pathology.

## Methods

### Literature search

A database search was conducted using MEDLINE, EMBASE and CINAHL (until December 2009). The search

strategy was designed to identify the publications for four separate diagnostic test accuracy reviews of an imaging technique i.e. MRI, CT, X-ray and myelography in the identification of lumbar spinal pathology. The search strategy was developed to find prospective or retrospective cohort or case-control studies assessing the diagnostic accuracy of MRI in the identification of lumbar spinal pathology (i.e. radicular syndrome, spinal stenosis, spinal tumours, spinal fractures, spinal infection/inflammation, disc herniation, spondylolisthesis, spondylolysis, ankylosing spondylitis, disc displacement, osteoporotic fractures and other degenerative disc diseases) in adult patients with LBP or sciatica.

MRI had to be compared to a reference test defined as (1) findings at surgery, (2) expert panel opinion or (3) diagnostic work up. Only published full reports with sufficient data to construct a diagnostic two-by-two table were included. Next to the electronic search, reference lists of all retrieved relevant publications were checked. Two reviewers (AV/RvR and MW) independently applied the selection criteria to all titles and abstracts and reviewed relevant full papers. Disagreements were resolved in a consensus meeting or by consulting a third review author (MvT) in case of persisting disagreement.

### Data extraction and methodological quality

Data extraction and quality assessment was performed by two review authors (RvR and MW) independently using a standardised form ensuring adequate reliability of collected data. Data were extracted on:

- Author, year of publication and journal;
- Study design
- Study population characteristics: pathology considered, age, gender, numbers of subjects for inclusion in study and analysis, clinical features with in- and exclusion criteria, level of measurement and setting
- Index and reference test characteristics: type of test, year and methods of execution, cut-off values, positivity thresholds and outcome scales
- Diagnostic parameters; diagnostic two-by-two table or parameters to reconstruct this table

Methodological quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) list consisting of 20 items to be scored as “yes”, “no” or “unclear” [19, 20]. Scoring criteria are available upon request. Disagreements were resolved by consensus. A radiologist (AG) was consulted for the assessment of the used test technology (item 13). No weights for different quality items or summary quality score were applied since the interpretation of summary scores is problematic and potentially misleading [21, 22].

## Statistical analysis

For each primary diagnostic study sensitivity and specificity with 95% confidence intervals (CI) were calculated and presented in a forest plot. Results were stratified into different pathology subgroups. For meta-analysis (for homogenous studies) a bivariate random effects analysis (using STATA 10 software) was used to calculate pooled estimates of sensitivity and specificity [23]. This method provides a random effects estimate of the mean pooled summary estimates of sensitivity and specificity with corresponding 95% confidence intervals thereby dealing with both within and between study variation together with any correlation that might exist between the sensitivity and the specificity. The resulting pooled summary estimates of sensitivity and specificity including a 95% confidence ellipse were plotted in ROC space and presented together with their corresponding prior probabilities, likelihood ratios (LR+ and LR–) and the diagnostic odds ratio (DOR). If no pooled estimates could be calculated, the range of sensitivity and specificity together with the prior probability was presented for each subgroup. Several sources may have contributed to heterogeneity in diagnostic accuracy parameters, which could only be addressed descriptively due to limited number of included studies.

## Results

### Literature search and data extraction

Figure 1 summarizes the search process, which resulted in eight articles on MRI being included in this review. In total these eight studies included 467 patients of which 1,476 discs or foramens were assessed. On average more males were included (50–78%). Six studies were prospectively designed. All studies were performed in a secondary care setting with most studies using surgery as the reference standard and one study using expert panel consensus. One study [27] presented accuracy data for both a full MRI and a limited MRI protocol performed in the same patients, only results obtained by the full MRI protocol were included since this protocol better resembled the protocols used in other studies including both axial and sagittal weighted images. The included studies were stratified according to the different pathologies: i.e. lumbar disc herniation [16, 24–29] and spinal stenosis [26, 30]. Lumbar disc herniation was subdivided into (1) Herniated nucleus pulposus (HNP) and (2) Nerve root compression due to HNP.

### Methodological quality

Figure 2 shows the scores for each quality item across the eight included studies, while the quality assessment of the

individual studies can be found in Fig. 3. Poor reporting of several quality items hindered assessment of the risk of bias and may have affected the validity of the reported sensitivities and specificities.

## Findings

The sensitivity and specificity of the studies are presented in a forest plot (Fig. 4), subdivided into the different pathology subgroups.

### Lumbar disc herniation

#### Herniated nucleus pulposus

Six studies [24–29] assessed the diagnostic accuracy of MRI for lumbar disc herniation i.e. HNP including sequestration, extrusion and protrusion or disc bulging. In all studies surgery was the reference standard. One study [25] presented the results for the combined identification of HNP and degenerative disc disease, consequently this study was not included in the pooled analysis.

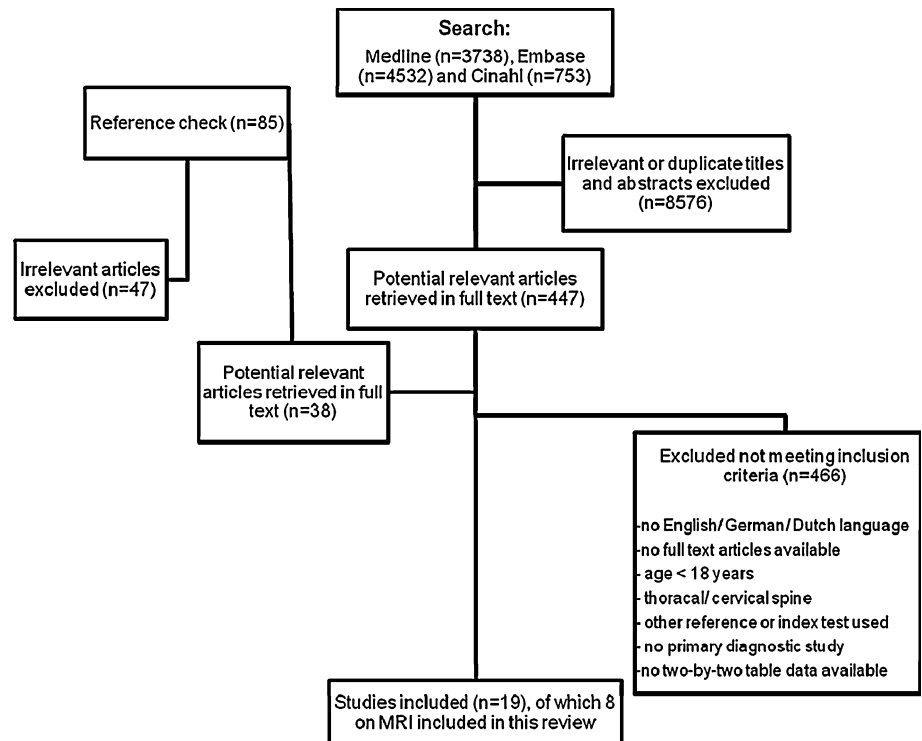
Five studies (197 patients; 322 discs) were considered sufficiently clinically homogenous for a pooled analysis. One study [27] assessed more than one disc level per patient. The prior probabilities ranged from 49% [26] to 77% [28] with a mean prior probability of 63%. The sensitivity of MRI for HNP ranged from 64 to 92% and the specificity ranged from 55% to even 100% [28]. The results of the bivariate analysis are plotted graphically in ROC space (Fig. 5). The pooled summary estimate of sensitivity and specificity were 75% (95% CI 65–83%) and 77% (95% CI 61–88%), respectively. This corresponds with a LR+ of 3.30 (95% CI 1.76–6.21), a LR– of 0.33 (95% CI 0.21–0.50) and a DOR of 10.12 (95% CI 3.88–26.39).

Due to small number of studies included, sources of heterogeneity were only explored descriptively. Both studies with a relatively low prior probability of 49% [26, 29] also presented slightly lower sensitivities (0.64 and 0.71 vs. 0.72–0.92) with similar specificities compared to those studies with a higher prior probability. Furthermore the two studies [27, 28] subjected to partial verification revealed a slightly higher sensitivity (0.83 and 0.92) compared to studies avoiding partial verification (0.64–0.72). There were no clear differences in specificity. The study [24] that did not use an appropriate MRI technology demonstrated a lower specificity yet a similar sensitivity when compared to those studies using an appropriate MRI technology.

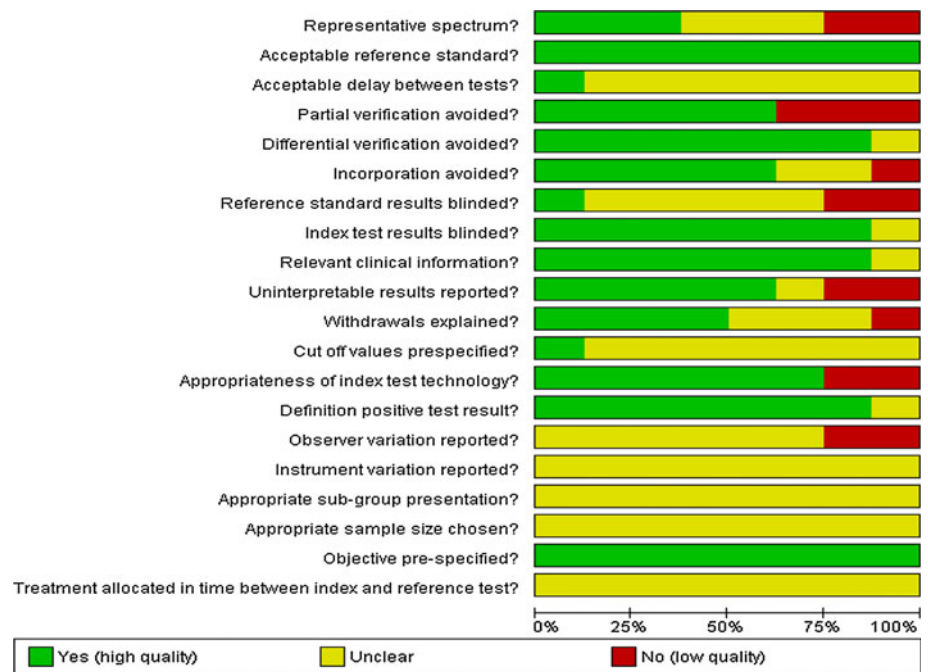
#### Nerve root compression due to HNP

The diagnostic accuracy of MRI for nerve root compression due to HNP was evaluated in two studies ( $n = 128$ )

**Fig. 1** Flow chart of literature search process



**Fig. 2** Results of the assessment of methodological quality items presented as percentage across all included studies



with prior probabilities of 77.9% [16] and 93.9% [27]. The results demonstrated a high sensitivity of 81 and 92% with varying specificity of 52 and 100%. The high specificity in the study of Chawalparit et al. might be due to partial

verification since not all patients underwent the reference test (surgery). The studies used a different reference standard i.e. findings at surgery versus expert panel consensus precluding statistical pooling.

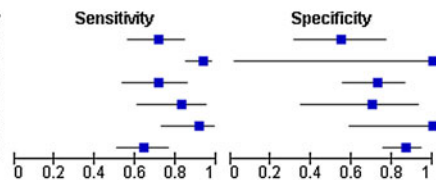
**Fig. 3** Methodological quality for each included study

	Representative spectrum?	Acceptable reference standard?	Acceptable delay between tests?	Partial verification avoided?	Differential verification avoided?	Incorporation avoided?	Reference standard results blinded?	Index test results blinded?	Relevant clinical information?	Uninterpretable results reported?	Withdrawals explained?	Cut off values prespecified?	Appropriateness of index test technology?	Definition positive test result?	Observer variation reported?	Instrument variation reported?	Appropriate sub-group presentation?	Appropriate sample size chosen?	Objective pre-specified?	Treatment allocated in time between index and reference test?	
Aota 2007	●	●	?	●	●	●	●	●	●	●	●	?	●	●	●	●	●	●	●	●	?
Bernard 1994	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Birney 1992	?	●	?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Bischoff 1993	?	●	?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Chawalparit 2006	●	●	?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Forristall 1988	●	●	?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Jackson 1989	?	●	?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Thornbury 1993	●	●	?	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

**Fig. 4** Forest plot of study results per pathology group describing sensitivities and specificities with accompanying 95% confidence intervals as well as the numbers of TP true positive, FP false positive, FN false negative and TN true negative results

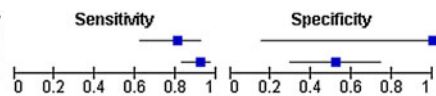
**Lumbar Disc Herniation**

Study	TP	FP	FN	TN	Sensitivity	Specificity
Bernard 1994	33	9	13	11	0.72 [0.57, 0.84]	0.55 [0.32, 0.77]
Birney 1992	70	0	5	1	0.93 [0.85, 0.98]	1.00 [0.03, 1.00]
Bischoff 1993	25	10	10	27	0.71 [0.54, 0.85]	0.73 [0.56, 0.86]
Chawalparit 2006	19	3	4	7	0.83 [0.61, 0.95]	0.70 [0.35, 0.93]
Forristall 1988	22	0	2	7	0.92 [0.73, 0.99]	1.00 [0.59, 1.00]
Jackson 1989	38	8	21	53	0.64 [0.51, 0.76]	0.87 [0.76, 0.94]



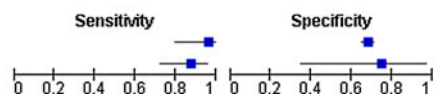
**HNP caused nerve root compression**

Study	TP	FP	FN	TN	Sensitivity	Specificity
Chawalparit 2006	25	0	6	2	0.81 [0.63, 0.93]	1.00 [0.16, 1.00]
Thornbury 1993	68	10	6	11	0.92 [0.83, 0.97]	0.52 [0.30, 0.74]



**Spinal Stenosis**

Study	TP	FP	FN	TN	Sensitivity	Specificity
Aota 2007	24	289	1	622	0.96 [0.80, 1.00]	0.68 [0.65, 0.71]
Bischoff 1993	34	2	5	6	0.87 [0.73, 0.96]	0.75 [0.35, 0.97]

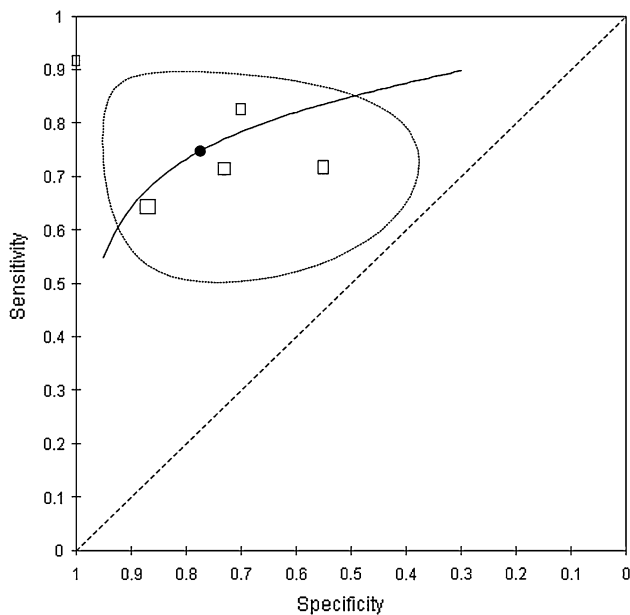


**Spinal stenosis**

Two studies described the accuracy of MRI in the identification of spinal stenosis [26, 30] in 983 foramina of 118 patients, and used surgery as the reference standard. The clinical homogenous studies demonstrated rather different prior probabilities (2.7% [30] and 83% [26]), possibly due to different population characteristics, or the unequal number of foramina/levels assessed in each study. Both studies showed high sensitivities of 87 and 96% coupled with lower specificities of 68 and 75%. Since this group only consisted of two studies, pooling of summary estimates was not performed.

**Discussion**

This review summarizes the evidence of the diagnostic accuracy of MRI in adult patients with LBP or sciatica in identifying specific lumbar spinal pathology. The pooled summary estimates for HNP sensitivity 75% (95% CI 65–83%) and specificity 77% (95% CI 61–88%) resulted in a positive predictive value of 84% and a negative predictive value of 64% given the mean prior probability of 63% in the included studies. This suggests that a substantial proportion of the patients will be incorrectly classified. The studies on spinal stenosis and nerve root compression due to HNP were not suitable for pooling. All results should be



**Fig. 5** Summary ROC plot presenting pooled estimates of sensitivity and specificity including 95% confidence ellipse and hierarchical ROC curve summary based on bivariate analysis of five studies describing the diagnostic accuracy of MRI for identifying lumbar disc herniation

interpreted cautiously, since the results were only based on a limited number of studies of moderate quality with several unaddressed sources of heterogeneity limiting the generalizability and validity of the results.

One of the most important sources of heterogeneity was the wide range in prior probabilities (from 2.7 to 82.9% for spinal stenosis [26, 30]). This is likely to influence the results as the posterior probability depends highly on the prior probability. The variation in prior probabilities will, at least partly, be the result of patient selection and setting; often scored as unclear or inadequate in the included studies. In this review studies with relatively high prior probabilities might be overrepresented as some studies [24, 28] may have included only those patients likely to undergo, or indeed underwent, surgery. Furthermore all studies were performed in a secondary care setting, where patients are likely to have a higher prior probability due to referral of only those patients with a relative high suspicion of specific pathology. Results only apply to these settings.

A second issue is the lack of a gold reference standard. This resulted in a large variation of reference tests. In this review, studies were only included if they used surgery, expert panel consensus or diagnostic work up as reference standard. Surgery, especially when combined with clinical follow-up, is often regarded as the best reference test, but subject to partial verification as often only patients with a strong suspicion of a specific underlying cause will be subjected to surgery. Verification bias might lead to a higher sensitivity and a lower specificity [31], yet it has

also been found that it increases both sensitivity and specificity [32]. Those studies [25, 27, 28] in which partial verification is likely to be present, indeed showed a somewhat higher sensitivity compared to studies without verification bias along with comparable or higher specificity.

A third issue is related to the index test characteristics, such as the reliability of the index test, i.e. the (inter- and intra-) observer variation, which may have affected the results due to lack of consensus on radiological definitions of findings related to subjective symptoms [33]. However, the extent of the effect is difficult to estimate, since the issue of observer variation was only addressed in two studies; both with relative poor scores ( $\kappa < 0.70$  for interobserver variation) [27, 30]. Most studies did not report objective cut-off values, used clear definitions or detailed procedures, possibly leading to different numbers of positive and negative test results. Furthermore, two of the included studies [24, 25] used older MRI techniques with less advanced visualising capacities probably resulting in a poorer identification of lumbar spinal pathology.

Finally, heterogeneity arose from the fact that some studies reported their results at disc level whereas others at patient level. Most studies assessed several discs per patient and some even presented unequal numbers of discs measured per patient, resulting in multiple (and unequal) inclusions of the same patient in the analyses. This leads to smaller confidence intervals and possibly to an overestimation of diagnostic accuracy. This occurs when patients with multiple signs of lumbar disc herniation at subsequent disc levels are more likely to be subjected to multiple testing than patients without these signs.

#### Strengths and weaknesses

To our knowledge this study is the first systematic review summarizing the available evidence on the diagnostic test accuracy of MRI for the identification of lumbar spinal pathology in LBP patients. A potential limitation of this review is the use of a study design filter in our search strategy. The rationale for the use of this filter was to limit the harvest of 18,239 citations found without the filter, even though using methodological filters has limitations [34]. To reduce the risk of missing important articles, the references of both included articles as well as review articles were checked, yet still there is always a chance of missing relevant publications.

Furthermore, in this review, the accuracy of the isolated use of MRI was assessed, though in routine clinical practice MRI will often be used in combination with other clinical observations or test results. Diagnosis and treatment decisions are never based on MRI findings only but always on the entire clinical assessment including MRI,

standing X-rays and clinical evaluation. Yet, our search strategy did not identify studies assessing the accuracy of MRI combined with other tests.

#### Implications for clinical practice and research

The importance of the findings should be interpreted in light of its clinical consequences. The value of MRI in identifying lumbar spinal pathology largely depends on the role of MRI results in clinical decisions regarding the management of LBP or sciatica and resulting outcomes. This could be either to exclude patients without the target condition to spare invasive treatments, or to identify as many patients as possible when delayed treatment results in worse patient outcomes. The role of MRI thereby largely depends on the suspected underlying pathology as well as the setting and patient characteristics.

This review demonstrates that a considerable proportion of patients will be incorrectly classified by MRI and may not be offered adequate management of LBP. However, the evidence for the diagnostic accuracy of MRI identified by this review is not conclusive and is limited to lumbar disc herniation and spinal stenosis. Consequently this cannot be generalized to other specific pathologies underlying LBP or sciatica. To provide more profound evidence on the role of MRI in the identification of lumbar spinal pathology in LBP patients, there is a strong need for high quality and accurately reported studies.

Additionally, in this review, the accuracy of the isolated use of MRI was assessed, although in routine clinical practice MRI will not be used isolated but in combination with other clinical observations or test results. Combined diagnostic information may provide different estimates of diagnostic accuracy that can be better placed in its clinical context. Future research should therefore address the added diagnostic value of MRI in the triage of LBP patients and its impact on subsequent treatment decisions.

**Acknowledgments** We thank Rene Otten for his useful help in designing the search strategy. This project was funded by Health Care Insurance Board (CvZ), Diemen, The Netherlands.

#### References

- Burton AK, Balague F, Cardon G et al (2006) Chapter 2. European guidelines for prevention in low back pain: November 2004. *Eur Spine J* 15(Suppl 2):S136–S168
- Ehrlich GE (2003) Low back pain 45. *Bull World Health Organ* 81:671–676
- Elders LA, Burdorf A (2004) Prevalence, incidence, and recurrence of low back pain in scaffolders during a 3-year follow-up study 42. *Spine (Phila Pa 1976.)* 29:E101–E106
- Chou R, Qaseem A, Snow V et al (2007) Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med* 147:478–491
- van TM, Becker A, Bekkering T et al (2006) Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 15(Suppl 2):S169–S191
- Kendrick D, Fielding K, Bentley E et al (2001) Radiography of the lumbar spine in primary care patients with low back pain: randomised controlled trial. *BMJ* 322:400–405
- Ash LM, Modic MT, Obuchowski NA et al (2008) Effects of diagnostic information, per se, on patient outcomes in acute radiculopathy and low back pain. *AJNR Am J Neuroradiol* 29:1098–1103
- Loblaw DA, Perry J, Chambers A et al (2005) Systematic review of the diagnosis and management of malignant extradural spinal cord compression: the Cancer Care Ontario Practice Guidelines Initiative's Neuro-Oncology Disease Site Group 13. *J Clin Oncol* 23:2028–2037
- Tsiodras S, Falagas ME (2006) Clinical assessment and medical treatment of spine infections. *Clin Orthop Relat Res* 444:38–50
- Jarvik JG, Deyo RA (2002) Diagnostic evaluation of low back pain with emphasis on imaging. *Ann Intern Med* 137:586–597
- Deyo RA, Bigos SJ, Maravilla KR (1989) Diagnostic imaging procedures for the lumbar spine. *Ann Intern Med* 111:865–867
- Ozturk C, Karadereler S, Ornek I et al (2010) The role of routine magnetic resonance imaging in the preoperative evaluation of adolescent idiopathic scoliosis. *Int Orthop* 34:543–546
- Chou R, Fu R, Carrino JA et al (2009) Imaging strategies for low-back pain: systematic review and meta-analysis 9. *Lancet* 373:463–472
- Jarvik JG, Hollingworth W, Martin B et al (2003) Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *JAMA* 289:2810–2818
- Lurie JD, Birkmeyer NJ, Weinstein JN (2003) Rates of advanced spinal imaging and spine surgery. *Spine (Phila Pa 1976.)* 28:616–620
- Thornbury JR, Fryback DG, Turski PA et al (1993) Disk-caused nerve compression in patients with acute low back pain: diagnosis with MR, CT myelography and plain CT. *Radiology* 186: 731–738
- Jensen MC, Brant-Zawadzki MN, Obuchowski N et al (1994) Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 331:69–73
- van Tulder MW, Assendelft WJ, Koes BW et al (1997) Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine (Phila Pa 1976.)* 22:427–434
- Smidt N, Deeks J, Moore T (2005) *Cochrane Handbook for Cochrane Reviews of Diagnostic Test Accuracy*. The Cochrane Collaboration. Ref Type: Generic
- Whiting P, Rutjes AW, Dinnes J et al (2004) Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 8:iii1–iii234
- Juni P, Witschi A, Bloch R et al (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 282: 1054–1060
- Whiting P, Harbord R, Kleijnen J (2005) No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol*. 5:19
- Reitsma JB, Glas AS, Rutjes AW et al (2005) Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 58:982–990
- Bernard TN Jr (1994) Using computed tomography/discography and enhanced magnetic resonance imaging to distinguish between scar tissue and recurrent lumbar disc herniation. *Spine (Phila Pa 1976.)* 19:2826–2832

25. Birney TJ, White JJ Jr, Berens D et al (1992) Comparison of MRI and discography in the diagnosis of lumbar degenerative disc disease. *J Spinal Disord* 5:417–423
26. Bischoff RJ, Rodriguez RP, Gupta K et al (1993) A comparison of computed tomography-myelography, magnetic resonance imaging, and myelography in the diagnosis of herniated nucleus pulposus and spinal stenosis. *J Spinal Disord* 6:289–295
27. Chawalparit O, Churojana A, Chiewvit P et al (2006) The limited protocol MRI in diagnosis of lumbar disc herniation. *J Med Assoc Thai* 89:182–189
28. Forristall RM, Marsh HO, Pay NT (1988) Magnetic resonance imaging and contrast CT of the lumbar spine. comparison of diagnostic methods and correlation with surgical findings. *Spine (Phila Pa 1976.)* 13:1049–1054
29. Jackson RP, Cain JE Jr, Jacobs RR et al (1989) The neuroradiographic diagnosis of lumbar herniated nucleus pulposus: II. a comparison of computed tomography (CT), myelography, CT-myelography, and magnetic resonance imaging. *Spine (Phila Pa 1976.)* 14:1362–1367
30. Aota Y, Niwa T, Yoshikawa K et al (2007) Magnetic resonance imaging and magnetic resonance myelography in the presurgical diagnosis of lumbar foraminal stenosis. *Spine (Phila Pa 1976.)* 32:896–903
31. Diamond GA (1992) Off Bayes: effect of verification bias on posterior probabilities calculated using Bayes' theorem 5. *Med Decis.Making* 12:22–31
32. Mol BW, Lijmer JG, van der MJ et al (1999) Effect of study design on the association between nuchal translucency measurement and Down syndrome. *Obstet Gynecol* 94:864–869
33. Jacobsen S, Gosvig KK, Sonne-Holm S (2006) Low back degeneration and radiology. *Ugeskr.Laeger* 168:769–773
34. Leeflang MM, Deeks JJ, Gatsonis C et al (2008) Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 149:889–897