Diagnostics in low back pain

Evelien de Schepper

Diagnostics in Low Back Pain

Evelien I.T. de Schepper

Department of General Practice
Erasmus MC, University Medical Center Rotterdam

Financial support for the publication of this thesis was kindly provided by the SBOH, employer of GP trainees, Anna Fonds Leiden, the MRI Center and the Erasmus University, Rotterdam, the Netherlands
ISBN: 978-94-6169-779-0
Cover: Marlou van Oosterhout, studio Moost, Rotterdam, the Netherlands
Layout: Marlou van Oosterhout, studio Moost en Optima Grafische Communicatie, Rotterdam, the Netherlands
Printing: Optima Grafische Communicatie, Rotterdam, the Netherlands

Diagnostics in Low Back Pain

Diagnostiek bij lage rugpijn

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

dinsdag 2 februari 2016 om 13.30

door

Evelien Irene Truus de Schepper
geboren te Bergen op Zoom

Erasmus University Rotterdam

Promotiecommissie

Promotor	Prof.dr. S.M.A. Bierma-Zeinstra
Overige leden	Prof.dr. Ir. A. Burdorf Prof.dr. G.P. Krestin Prof.dr. W.C. Peul
Copromotor	Dr. P.A.J. Luijsterburg

Reumafonds
Dutch Arthritis Foundation

The research described in this thesis was supported by a grant of the Dutch Arthritis Foundation

Contents

1 General introduction

Low back pain is a major health problem. A publication from 2013 showed that, from the 289 diseases and conditions that were investigated, low back pain leads to the most 'years lived with disability' [1]. It is one of the most frequently occurring musculoskeletal disorders and leads to high costs due to disability, lost time from work and medical care. The point prevalence of back pain in the general population ranges from 12 to 30% [2, 3]. Several studies found that between 30% and 45% of those with back pain have had contacts with their general practitioner within a period of one year, which means that the general practitioner is a healthcare provider that is often visited for back pain [4-6]. In the Netherlands, the incidence of low back pain in general practice is about 30 episodes per 1,000 patients registered per year [3, 7].

Low back pain is usually divided into 'specific' and 'non-specific' low back pain. Specific low back pain is defined as symptoms caused by a particular pathophysiological mechanism, such as hernia nuclei pulposi, discitis, malignancy and fracture. Non-specific low back pain is defined as pain for which no specific cause can be shown; in about 80 to 95% of people with low back pain, no specific cause is found [8].

Adequate treatment of patients with low back pain begins with making the correct diagnosis. In the Netherlands, the first step in diagnostics for patients with low back pain is mostly made by the general practitioner, almost exclusively based on history and physical examination. In addition, in a small proportion of the patients, diagnostic imaging is needed. International guidelines recommend the use of imaging only when there is suspicion of serious pathology (fracture, malignancy and discitis), or in patients with severe sciatica for whom surgery is indicated because they fail to respond to conservative care for at least 6 to 8 weeks [9, 10]. Currently, there is widespread consensus that there is no indication to perform diagnostic imaging in patients with non-specific low back pain.
But despite of all these recommendations, there is still little consensus, either within or between specialties, on appropriate diagnostic evaluation of low back pain [11]. Lumbar spine imaging (plain radiography, CT, and MRI) is still often performed in patients with low back pain, and it is often performed in the absence of a clear indication for it [9, 12]. These findings support the hypothesis that much of the variation in medical care for low back pain may be due to physician uncertainty [11].

The aim of the work presented in this thesis is to gain more insight into the diagnostic tools used in patients with low back pain. First, we examined how lumbar MRI is currently used by general practitioners in the Netherlands.

MRI in primary care

If a physician suspects the presence of a specific disease, diagnostic imaging can be used. In recent years, general practitioners in the Netherlands can refer low back pain patients for MRI of the lumbar spine themselves. Possible reasons for the use of MRI in general practice are i) to detect or exclude specific pathologies, ii) to reassure the patient (and physician), and/or iii) to prevent unnecessary referrals to secondary care. Despite the recommendations of the guidelines to use MRI only in specific cases, the use of MRI as the initial imaging for low back pain seems to have become more common in general practice in countries such as the USA and Australia [9, 13]. However, data on the use of MRI by general practitioners in the Netherlands are still lacking. Therefore, we designed an observational prospective cohort study with a 12-month follow-up in patients with low back pain referred for MRI in general practice. In **Chapter 2**, we explore the characteristics and MRI findings of patients with low back pain referred for MRI in general practice.

The ultimate goal of any diagnostic test is to improve the clinical outcome of the patient. Well-conducted randomized trials are the top of the diagnostic evidence hierarchy, because they provide the most direct information on the clinical benefits and harms of alternative testing strategies. However, in daily practice most studies on diagnostic tests estimate how accurately they can identify a disease or condition, or how well the test provides prognostic information. Furthermore, identification of prognostic factors predicting recovery, persistent pain, and disability are important for better understanding of the clinical course, to inform patients and physicians, and to support therapeutic decision making [14]. In **Chapter 3**, we investigate the added prognostic value of baseline MRI findings over known prognostic factors for patient recovery.

The use of MRI by general practitioners to diagnose and manage patients with low back pain might result in avoidance of unnecessary hospital referrals. However, data on the patterns of subsequent care among patients

referred for lumbar spine MRI by their general practitioner are scarce. Better understanding of the patterns of healthcare services used after MRI of the lumbar spine will provide information on how MRI findings are used by general practitioners for subsequent management. Identification of possible prognostic factors predicting consultation with specialists or surgery, can be important to inform patients and physicians. Therefore, in **Chapter 4**, we investigate the association between patient characteristics, back pain characteristics and MRI abnormalities with subsequent specialist consultation and/or surgery in the same low back pain patients referred to MRI by their general practitioner.

Lumbar spinal stenosis

One of the anatomical abnormalities that can be found with MR imaging is lumbar spinal stenosis. Lumbar spinal stenosis is commonly used to describe patients with symptoms related to an anatomic reduction of the lumbar spinal canal size [15]. The challenge to the anatomically-based definition is that, while necessary for the diagnosis of lumbar spinal stenosis, it is not sufficient to determine the severity of symptoms that leads a patient to seek treatment [15]. The extent of narrowing of the spinal canal correlates poorly with symptom severity, and radiological significant lumbar stenosis can also be found in asymptomatic individuals [15-18]. Furthermore, lower extremity pain, and numbness or weakness are frequently seen in the setting of low back pain, and many other causes may also be involved. Therefore, correlating symptoms and physical examination findings with imaging results is necessary to establish a definitive diagnosis. A wide range of clinical, electrodiagnostic and radiological tests are currently used to diagnose lumbar spinal stenosis. It is important to know the diagnostic value of these tests because false-positive test results may lead to unnecessary surgery and/or costly or invasive additional diagnostic interventions. In **Chapter 5**, we update our earlier systematic review on the diagnostic accuracy of tests used to diagnose lumbar spinal stenosis.

Lumbar disc degeneration

Once patients with radiculopathy and serious causes of back pain (such as fracture and malignancy) are excluded, the remaining patients (approximately 85%) are those with so-called ‘non-specific low back pain’.

In these patients clinicians apply generic symptomatic treatments such as advice to stay active and avoid bed rest, as well as analgesic medicines, exercise and manipulation. While this approach is simple, unfortunately it does not always work very well. The limitations of current approaches are illustrated by many systematic reviews of treatments for low back pain that reveal that existing treatments for non-specific low back pain have, at best, only small effects [19-21]. One reason for this might be that the 'one-size fits all approach' advocated by guidelines fails to target treatments at patients who might benefit the most, thus diluting their potential benefits [22]. Identifying subgroups of patients for whom different treatments are superior has been referred to as the 'Holy Grail' of low back pain research. One of the possible subgroups based on pathoanatomical findings are patients with degeneration of the spine. In the second part of this thesis, we performed an epidemiological study on the characteristics of lumbar disc degeneration.

Lumbar disc degeneration is characterized (radiographically) by the presence of osteophytes, endplate sclerosis, and disc space narrowing. The association between low back pain and degenerative changes in the lumbar spine is complex. However, it is known that lumbar disc degeneration can be a possible risk factor for back pain in adults, with odds ratios ranging from 1.3 to 3.2 [23, 24]. Nevertheless, due to differences in the definitions used for lumbar disc degeneration, studies on the relation between low back pain and lumbar disc degeneration are difficult to compare. This is why it is still difficult to provide a clinically relevant definition for lumbar disc degeneration. Therefore, the study in **Chapter 6** explores the association between various individual radiographic features of lumbar disc degeneration and self-reported low back pain.

Patients with low back pain with symptoms due to lumbar disc degeneration could be a subgroup within the population of non-specific low back pain patients. Clinical symptoms associated with radiographic lumbar disc degeneration may help identify such patients. Although lumbar disc degeneration cannot be defined as 'real' osteoarthritis because the facet joints are the only synovial joints in the spine, it is often used as a proxy for osteoarthritis of the spine. Clinical osteoarthritis of the knee and hip includes (besides pain) also morning stiffness (ACR criteria) [25, 26]. Therefore, in **Chapter 7**, we examine the association between spinal morning stiffness and lumbar disc degeneration. These associations are

also compared with the associations between morning stiffness in the legs, and knee or hip osteoarthritis.

In **Chapter 8** we studied the possible association between osteoarthritis of the spine and hip pain. Differentiating back pain from hip pain in patients who present with classic signs and symptoms is mostly not difficult. However, in some cases, patients present with nonspecific complaints of pain in the buttock, lateral hip or thigh. The differential diagnosis of this nonspecific pain is broad and includes radiating pain from the lumbar spine. The differentiation of signs and symptoms suggestive of hip disorders versus spine disorders is important in giving patients the most beneficial treatment, especially if the treatment includes a major reconstructive surgery, such as hip replacement. Preoperative identification of factors associated with hip pain arising from the lumbar spine would help physicians by identifying the subgroup of patients who might not experience full relief of pain with a hip arthroplasty. Therefore, the study in Chapter 8 explores the association between self-reported hip pain and various individual radiographic features of spinal osteoarthritis by vertebral level, including osteophytes and disc space narrowing.

Finally, in **Chapter 9** the key findings of the previous chapters are summarized and discussed in the context of current knowledge and evidence, and directions for future research are provided.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380: 2163-2196.
2. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. *Arthritis Rheum* 2012; 64: 2028-2037.
3. M.W. van Tulder, B.W. Koes. Evidence-based handelen bij lage rugpijn 2013 Bohn Staflue van Loghum, Houten.
4. Andersson HI, Ejlertsson G, Leden I, Schersten B. Impact of chronic pain on health care seeking, self care, and medication. Results from a population-based Swedish study. *J Epidemiol Community Health* 1999; 53: 503-509.
5. Molano SM, Burdorf A, Elders LA. Factors associated with medical care-seeking due to low-back pain in scaffolders. *Am J Ind Med* 2001; 40: 275-281.
6. Waxman R, Tennant A, Helliwell P. Community survey of factors associated with consultation for low back pain. *BMJ* 1998; 317: 1564-1567.
7. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 2003; 102: 167-178.
8. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA* 1992; 268: 760-765.
9. Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am* 2012; 50: 569-585.
10. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ* 2007; 334: 1313-1317.
11. Cherkin DC, Deyo RA, Wheeler K, Ciol MA. Physician variation in diagnostic testing for low back pain. Who you see is what you get. *Arthritis Rheum* 1994; 37: 15-22.
12. Pham HH, Landon BE, Reschovsky JD, Wu B, Schrag D. Rapidity and Modality of Imaging for Acute Low Back Pain in Elderly Patients. *Archives of Internal Medicine* 2009; 169: 972-981.
13. Williams CM, Maher CG, Hancock MJ, McAuley JH, McLachlan AJ, Britt H, et al. Low Back Pain and Best Practice Care A Survey of General Practice Physicians. *Archives of Internal Medicine* 2010; 170: 271-277.
14. Verwoerd AJ, Luijsterburg PA, Lin CW, Jacobs WC, Koes BW, Verhagen AP. Systematic review of prognostic factors predicting outcome in non-surgically treated patients with sciatica. *Eur J Pain* 2013; 17: 1126-1137.

15. Genevay S, Atlas SJ. Lumbar spinal stenosis. *Best Pract Res Clin Rheumatol* 2010; 24: 253-265.
16. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990; 72: 403-408.
17. Kalichman L, Cole R, Kim DH, Li L, Suri P, Guermazi A, et al. Spinal stenosis prevalence and association with symptoms: the Framingham Study. *Spine J* 2009; 9: 545-550.
18. Wiesel SW, Tsourmas N, Feffer HL, Citrin CM, Patronas N. A study of computer-assisted tomography. I. The incidence of positive CAT scans in an asymptomatic group of patients. *Spine (Phila Pa 1976)* 1984; 9: 549-551.
19. Hancock MJ, Maher CG, Laslett M, Hay E, Koes B. Discussion paper: what happened to the 'bio' in the bio-psycho-social model of low back pain? *Eur Spine J* 2011; 20: 2105-2110.
20. Deyo RA. Treatments for back pain: can we get past trivial effects? *Ann Intern Med* 2004; 141: 957-958.
21. Machado LA, Kamper SJ, Herbert RD, Maher CG, McAuley JH. Analgesic effects of treatments for non-specific low back pain: a meta-analysis of placebo-controlled randomized trials. *Rheumatology (Oxford)* 2009; 48: 520-527.
22. Foster NE, Hill JC, Hay EM. Subgrouping patients with low back pain in primary care: are we getting any better at it? *Man Ther* 2011; 16: 3-8.
23. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine (Phila Pa 1976)* 1997; 22: 427-434.
24. Raastad J, Reiman M, Coeytaux R, Ledbetter L, Goode AP. The Association Between Low Back Pain and Radiographic Features: A Systematic Review with Meta-Analysis. *Arthritis & Rheumatology* 2014; 66: S78-S79.
25. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991; 34: 505-514.
26. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986; 29: 1039-1049.

2

Prevalence of spinal pathology in back pain patients referred for MRI in primary care

Evelien I.T. de Schepper, Bart W. Koes, Erik F.H. Veldhuizen, Edwin H.G. Oei, Sita M.A. Bierma-Zeinstra, Pim A.J. Luijsterburg

Abstract

Design

Cross-sectional, observational cohort study.

Background

The use of MRI as the initial imaging for low back pain has increased in general practice. However, because few data are available on the characteristics of these referred patients, more insight will provide information on how MRI scans are being used by general practitioners.

Objective

To describe the characteristics and MRI findings of low back pain patients from general practice referred for MRI, and investigate whether baseline characteristics differ between patients with and without specific findings seen on MRI.

Methods

Patients referred by their general practitioner for MRI of the lumbar spine were recruited. The MRI radiology reports were scored for the presence of disc bulging, disc herniation, nerve root compression, spinal stenosis, spondylolisthesis and serious pathologies (i.e. fracture, malignancy, discitis). Information on patients' characteristics, characteristics of the complaints, and red flags were derived from the baseline measurement. Cross-sectional differences between patients with and without specific MRI findings were analyzed using a Mann-Whitney U-test or a Fisher's exact test.

Results

A total of 683 low back pain patients were included; mean age was 49.9 (range 19-80) years and 53% was male. Mean back pain severity score was 6.6 (SD 2.0) and 67% of the patients reported having chronic low back pain. Of all MRI reports, 69% mentioned signs of nerve root compression. Serious pathologies were reported in 3% of the patients. Patients with malignancy were older and less often reported a history of back pain complaints (40% vs. 81%).

Conclusions

Most of these low back pain patients referred for MRI in general practice reported long-lasting complaints and severe low back pain. The MRI reports showed a relatively high number of serious spine pathologies.

Introduction

In the Netherlands, general practitioners (GPs) can refer low back pain patients for magnetic resonance imaging (MRI) of the lumbar spine. Possible reasons for the use of MRI in general practice are to detect or exclude specific pathologies, to reassure the patient (and physician), and/or to prevent unnecessary referrals to secondary care.
Clinical practice guidelines recommend not to immediately initiate lumbar spine imaging in patients with acute or subacute low back pain and without features (red flags) suggesting serious underlying conditions [1-4]. Despite this recommendation, the use of MRI as the initial imaging for low back pain has become more common not only in secondary care, but also in primary care [1, 5]. In response, recommendations have been made to increase efforts to curb the overuse of lumbar MRI [1]. Nevertheless, a more widespread use of MRI in general practice is also advocated by others [6].
Meanwhile, data on the characteristics of low back pain patients referred to MRI by their GP are lacking. Therefore, to gain insight into the value of MRI in primary care, more information is needed on low back pain patients referred to MRI by their GP and how MRI scans are used by GPs. Furthermore, there is uncertainty regarding the prevalence of serious pathologies in primary care as an underlying cause of low back pain [7, 8]. Henschke et al. reported that serious pathologies in patients presenting to a primary care provider with acute low back pain are rare [9]. Although the prevalence of serious pathology in referred patients is expected to be higher, clear evidence is lacking.
The aims of this study are to describe the characteristics and MRI findings of patients with low back pain from general practice referred for MRI, and investigate whether baseline characteristics differ between patients with and without specific findings seen on MRI.

Methods

This study used the baseline data of a prospective observational cohort study in general practice, with a 12-month follow-up. Eligible patients were enrolled between June 2010 and September 2011. The study protocol was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, the Netherlands.

Study population

Consecutive eligible adults referred by their GP for MRI of the lumbar spine were recruited at the MRI Center (the Netherlands). The inclusion criteria for the study were aged ≥ 18 years and being referred by their GP for MRI of the lumbar spine. Patients were excluded if there were contraindications for undergoing MRI, or if the patient had insufficient understanding of the Dutch language and/or was incapable of understanding the ramifications of participation.

Eligible patients received written information about the study when they made an appointment at the MRI Center and could ask any additional questions up until the date of the MRI scan. Patients that declared interest in the study provided informed consent.

MRI findings

All patients underwent MRI (1.5 Tesla, Siemens, Erlangen, Germany) as scheduled. The MRI protocol consisted of sagittal and transverse T1 and T2-weighted sequences. We performed transverse imaging and a three-dimensional (3D) steady-state sequence (CISS) through affected disks and vertebrae. The MRIs were assessed by one of 7 radiologists of the MRI Center. As this study was designed to reflect routine general practice as closely as possible, we scored the findings described in the MRI radiology reports retrieved from the MRI Center, which were identical copies of the reports sent to the referring GP. There was no interference with the care given by the GP or any other healthcare providers with respect to advice, diagnostics or treatment.

A single trained reader [EdS], who was blinded to the participants' clinical data, extracted data from the MRI reports regarding the presence or absence of the following findings at each lumbar level (T12-L1 through L5-S1): intervertebral disc bulging, disc herniation (protrusion/extrusion per level), disc sequestration, nerve root compression, spinal stenosis, spondylolisthesis, and serious pathology (fracture, malignancy and/or discitis).

Questionnaires

After inclusion, the baseline measurement included validated questionnaires, for which participants were invited by email containing a secured link to the online questionnaires. The follow-up period was 12 months, with follow-up measurements at 3 and at 12 months. Reminders were sent by email after 2 and 3 weeks of non-response.

The baseline questionnaire included questions on: 1) patient characteristics: age, gender, body mass index (BMI), level of education (seven categories, ranging from 'primary school' to 'graduate school'), employment status, attitude and beliefs about low back pain at baseline measured with the Back Beliefs Questionnaire (BBQ, range 9-45) [10] in which a higher score indicates better attitude/belief regarding back pain, and quality of life (EuroQol, range -0.329 to 1.0 [11]; 2) back pain characteristics: history of back pain, duration of back symptoms (five categories, ranging from 'less than 6 weeks' to 'more than one year'), severity of the low back pain at baseline measured on an 11-point numerical rating scale (NRS) [12] in which 0 represents 'no pain' and 10 represents 'the worst pain ever', presence of radiating pain in the legs below the knee, severity of radiating pain in the leg (NRS), neurological symptoms of the legs determined with the question: "Did you have any complaints of numbness or tingling of the leg(s), and/or weakness of the leg(s) during the last week" (answer yes/no), morning stiffness of the back, disability measured with the Roland Disability Questionnaire (RDQ, range 0-24) [13], and history of back surgery; and 3) red flags: derived from the Dutch clinical practice guideline for low back pain [14] for the presence of two of the most common serious pathologies (vertebral fracture and malignancy) [9]. The red flags studied for the presence of vertebral fracture included: trauma, age ≥ 70 years, and female gender. The red flags studied for the presence of malignancy included: back pain started after age 50 years, presence of continuous back pain independent of posture or activity, night-time back pain, history of malignancy, and unexplained weight loss.

Statistical analysis

Data were checked for inconsistencies and missing baseline data were imputed using multivariate imputation resulting in five imputed datasets [15]. To enable easy interpretation, the following categorical variables were dichotomized: education was dichotomized in low level (compulsory education or lower secondary school) and high level; and duration of back pain in acute (less than 3 months) and chronic back pain (3 months and over).

Descriptive analyses were used to report the baseline characteristics of the patients using the mean and standard deviation (SD) for continuous data, and proportions for categorical data. Secondly, patients without any MRI findings were compared with those with MRI findings, and patients with a specific finding seen on MRI (disc herniation with nerve root compression,

spinal stenosis, spondylolisthesis, fracture, malignancy or discitis) were compared with those without this specific finding.
For the analyses, the Fisher's exact test was used to compare the categorical variables and a Mann-Whitney U-test was used to compare the numeric variables. All analyses were performed using SPSS version 20.0 (SPSS Inc., USA).

Results

A total of 683 patients participated in the study (Figure 1). At baseline, information on BMI was missing for 8 patients (1%) and information on severity of leg pain was missing for 2 patients (0.3%). The baseline characteristics of the patients are presented in Table 1; mean age was 49.9 years (SD 12.5; range 19-80) years. In total, 53% of the patients were male, 36% reported a low education level and 70% had a paid job.
Mean back pain severity score was 6.6 (SD 2.0) on a 0-10 scale, mean disability score on the RDQ was 13.5 (SD 5.2), and 67% of the patients

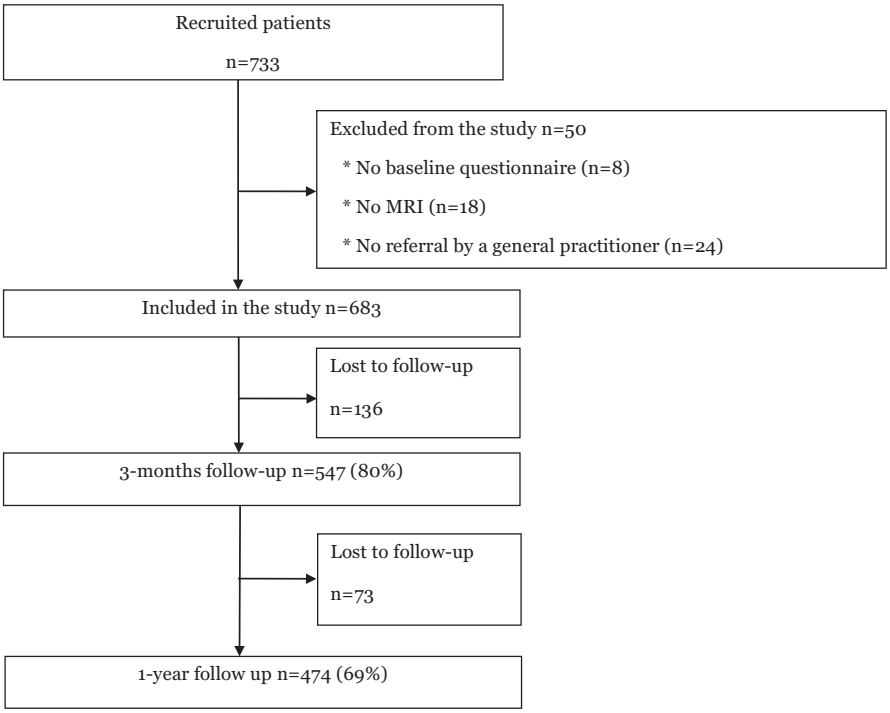

Figure 1. Flow chart of the study population

Table 1 Baseline characteristics of 683 included patients presenting for lumbar MRI as referred by their GP

	Study population *n*=683	**No abnormal findings *n*=44**	**Disc herniation with nerve root compression *n*=397**	**Spinal stenosis *n*=87**	**Spondylolisthesis *n*=56**	**Fracture *n*=17**	**Malignancy *n*=5**	**Discitis *n*=2**
Patient characteristics								
Age in years, mean (SD)	49.9 (12.5)	47.5 (14.1)	49.1 (12.4)*	55.9 (13.7)**	58.1 (12.7)**	56.4 (14.6)	66.1 (9.7)**	59.3 (2.3)
Male	365 (53)	13 (30)**	251 (63)**	58 (67)**	31 (55)	7 (41)	3 (60)	1 (50)
BMI, mean (SD)	25.9 (3.8)	24.1 (4.1)**	26.1 (3.5)*	26.7 (4.0)	25.9 (4.0)	25.9 (4.0)	24.4 (4.8)	28.5 (3.4)
Education low	244 (36)	15 (34)	139 (35)	41 (47)*	18 (32)	7 (41)	2 (40)	1 (50)
Employed (paid job)	479 (70)	33 (75)	290 (73)	45 (52)**	23 (41)**	10 (59)	2 (40)	1 (50)
Attitude and beliefs about back pain (BBQ), mean (SD)	26.3 (6.1)	27.0 (6.5)	26.5 (6.2)	26.1 (6.7)	26.0 (6.1)	25.6 (7.4)	30.0 (6.2)	19.5 (3.5)
EuroQol, mean (SD)	0.48 (0.3)	0.47 (0.3)	0.47 (0.3)	0.48 (0.3)	0.51 (0.3)	0.55 (0.3)	0.54 (0.4)	0.22 (0.1)
Back pain characteristics								
History of back pain	549 (80)	31 (71)	323 (81)	64 (74)	41 (73)	14 (82)	2 (40)	1 (50)
Chronic back pain > 3 months	455 (67)	34 (77)	240 (60)**	58 (67)	39 (70)	14 (82)	2 (40)	2 (100)
Severity of back pain (NRS), mean (SD)	6.6 (2.0)	7.0 (1.9)	6.4 (2.2)	6.5 (2.0)	6.5 (2.0)	6.8 (1.7)	7.2 (0.8)	4.5 (6.4)
Pain radiating in the leg below the knee	450 (66)	21 (48)*	297 (75)**	64 (74)	44 (79)*	11 (65)	4 (80)	1 (50)
Severity of leg pain (NRS), mean (SD)	5.6 (2.8)	4.7 (3.4)	6.1 (2.6)**	6.1 (2.6)	5.8 (2.5)	5.4 (2.7)	6.0 (2.3)	7.5 (0.7)
Neurological symptoms legs	525 (77)	28 (64)*	325 (82)**	69 (79)	45 (80)	11 (65)	3 (60)	2 (100)

Table 1 Baseline characteristics of 683 included patients presenting for lumbar MRI as referred by their GP (continued)

	Study population *n*=683	**No abnormal findings *n*=44**	**Disc herniation with nerve root compression *n*=397**	**Spinal stenosis *n*=87**	**Spondylolisthesis *n*=56**	**Fracture *n*=17**	**Malignancy *n*=5**	**Discitis *n*=2**
Morning stiffness of the back	353 (52)	26 (59)	199 (50)	40 (46)	31 (55)	10 (59)	3 (60)	1 (50)
Disability (RDQ), mean (SD)	13.5 (5.2)	13.3 (5.3)	14.0 (4.9) **	13.9 (4.9)	13.0 (5.2)	12.6 (5.4)	14.8 (5.6)	18.0 (0.0)
History of back surgery	112 (16)	10 (23)	63 (16)	19 (22)	14 (25)	4 (24)	1 (20)	0 (0)

Data are presented as numbers (percentages) unless otherwise indicated

*A significant difference ($p<0.05$) was found between patients with and without the specific finding as seen on MRI

**A significant difference ($p<0.01$) was found between patients with and without the specific finding as seen on MRI

MRI: Magnetic Resonance Imaging; GP: general practitioner; SD: standard deviation; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; EuroQol: higher scores represent higher quality of life (range -0.329-1.0); NRS: numeric rating scale (range 0-10, 0 means no pain); RDS: Roland disability questionnaire (range 0-24), a higher score indicates worse health

reported having chronic back pain. Of the 455 patients with chronic back pain at baseline, 324 patients (71%) reported back pain persisting for ≥ 1 year. In total, 66% of the patients reported pain radiating to the leg below the knee, mean leg pain severity score was 5.6 (SD 2.8), and 77% of the patients reported neurological symptoms of the leg(s).

MRI findings

An absence of abnormal MRI findings was reported in only 6% of the patients (Table 2). The MRI reports described disc herniation in 72% of the

Table 2 MRI findings of the 683 included patients

	Patients *n*=683	
No abnormal MRI findings	**44**	**(6)**
Disc bulging	**308**	**(45)**
Single level	103	(33)
Multiple levels	205	(67)
Disc herniation	**492**	**(72)**
Single level	265	(54)
Multiple levels	227	(46)
Sequester	42	(9)
T12-L1	10	(2)
L1-L2	37	(8)
L2-L3	66	(13)
L3-L4	101	(21)
L4-L5	285	(58)
L5-S1	272	(55)
Nerve root compression	**472**	**(69)**
Spinal stenosis	**87**	**(13)**
Spondylolisthesis	**56**	**(8)**
Grade I	42	(75)
Grade II	5	(9)
Grade III	0	(0)
Grade IV	0	(0)
Unknown	9	(16)
Impression fracture	**17**	**(3)**
Recent fracture	4	(24)
Malignancy	**5**	**(0.7)**
Discitis	**2**	**(0.3)**

Data are presented as numbers (percentages)
MRI: Magnetic Resonance Imaging

patients and 46% of these patients had disc herniation at multiple levels. In terms of their distribution by vertebral level, disc herniation was more frequent at the lower lumbar disc levels. In total, 69% of the MRI reports mentioned signs of nerve root compression. Disc herniation with nerve root compression was reported in 397 (58%) patients.
Spinal stenosis was reported in 87 patients (13%). The MRI reports described spondylolisthesis in 56 patients (18%) and 75% of these patients had a spondylolisthesis grade I. Regarding serious pathologies, 17 MRI reports (3%) mentioned fracture(s), and 4 of these (24%) mentioned a recent fracture of which 2 were caused by metastatic cancer. Malignancy was reported in 5 patients (0.7%) and discitis was reported in 2 patients (0.3%).

Patients without any abnormal MRI findings (6%) were more often female, had a slightly higher BMI, and less often reported pain radiating to the leg below the knee and neurological symptoms of the legs, compared with those with specific MRI findings (Table 1).
Comparison of patients with and without specific MRI findings also revealed a number of significant differences. Patients with disc herniation with nerve root compression were younger, more often male, had a slightly higher BMI, and more often reported acute back pain compared with those without. Also, they more often reported pain radiating to the leg below the knee and neurological symptoms of the leg(s). In addition, they had more severe leg pain (mean 6.0) and a higher level of disability as measured with the RDQ (mean 14.0).
Patients with spinal stenosis were more often male and of higher age compared to those without. Also, they more often reported a low education level and unemployment; however, these differences might be attributed to their higher age. Patients with spondylolisthesis more often reported pain radiating to the leg below the knee and were of higher age compared with those without. Again, the high level of unemployment might be attributed to their higher age.

Serious pathology
Patients with an impression fracture were of higher age (borderline significant: p=0.08) compared to those without. Patients with malignancy were significantly older (mean 65.2 years) and less often reported a history of back pain complaints (borderline significant; p=0.05) (40%).

The red flag 'trauma' was reported in 6 patients (35%) with an impression fracture; the red flag 'age over 70 years' was reported in 4 patients (24%)' and the red flag 'female gender' was reported in 10 patients (59%) (Table 3). In total, 41% of the patients with fracture reported 1 red flag, 29% of the reported 2 red flags, and 1 patient (6%) reported 3 red flags.
All 5 patients with malignancy reported the red flag 'back pain started after age 50 years'. None of the patients with malignancy reported the presence of continuous back pain independent of posture or activity. In total, 2 patients (40%) with a malignancy reported 1 of the five red flags, and 3

Table 3 Frequency of red flag signs and symptoms reported by all included 683 patients, 17 patients with fracture and 5 patients with malignancy seen on MRI.

	Patients ***n*=683**		**Fracture** ***n*=17**		**Malignancy** ***n*=5**	
Red flags for fracture						
Trauma	86	(13)	6	(35)	-	
Age over 70 years	30	(4)	4	(24)	-	
Female gender	318	(47)	10	(59)	-	
Red flags for malignancy						
Age at onset over 50 years	161	(24)	-		5	(100)
Continuous back pain	347	(51)	-		0	(0)
Back pain at night	364	(53)	-		2	(40)
History of malignancy	22	(3)	-		1	(20)
Unexplained weight loss	79	(12)	-		0	(0)
Red flags fracture						
No positive red flags	310	(45)	4	(24)	-	
1 positive red flag	316	(46)	7	(41)	-	
2 positive red flags	53	(8)	5	(29)	-	
3 positive red flags	4	(1)	1	(6)	-	
Red flags malignancy						
No positive red flags	127	(19)	-		0	(0)
1 positive red flag	235	(34)	-		2	(40)
2 positive red flags	233	(34)	-		3	(60)
3 positive red flags	80	(12)	-		0	(0)
4 positive red flags	8	(1)	-		0	(0)
5 positive red flags	0	(0)	-		0	(0)

Data are presented as numbers (percentages).
Red flags refer to signs and symptoms which indicate urgent need for treatment, or indicate a severe condition.

patients (60%) reported 2 red flags. Thus, all patients with malignancy reported at least 1 red flag at baseline.

Discussion

This study presents the characteristics and MRI findings of low back pain patients from general practice referred for MRI of the lumbar spine. A total of 683 low back pain patients were included; mean back pain severity score was 6.6 (SD 2.0), mean leg pain severity score was 5.6 (SD 2.8), and 67% of the patients had chronic low back pain. In total, 69% of the MRI reports mentioned signs of nerve root compression. Serious pathologies were reported in 3% and patients with malignancy were older and less often reported a history of back pain complaints.

In this study, most of the referred patients reported long-lasting back complaints. The baseline back pain severity and disability scores were generally higher than reported in earlier cohort studies on low back pain [16-18], and in three studies on low back pain patients referred for MRI or radiography by their primary physician [19-21]. Compared with the study of Jarvik et al. [20], our patients also more often reported pain radiating to the leg below the knee and more often a history of back surgery. A possible reason for these differences might be the Dutch healthcare system. In the Netherlands, GPs have a gatekeeping role in the healthcare system, i.e. patients need a referral from their GP to consult a hospital specialist. To maintain that gatekeeping role, GPs tend to adhere to the Dutch guidelines [14]. These guidelines recommend that diagnostic imaging should only be performed in patients who have chronic, severe radiculopathy, or with features suggesting a serious or specific underlying condition. Therefore, the patients in the present study may be worse off in terms of pain, disability and duration of complaints.

The MRI reports showed disc herniation in 72% and nerve root compression in 69% of the patients. These prevalences are much higher than reported in persons without low back pain [22]. They are also higher than reported in cohorts of patients with chronic low back pain [23] and acute back pain [24], and in cohorts including patients referred for radiography [20] and MRI [25]. Our prevalences match those reported by Hancock et al. [26] who compared rates of MRI findings of 30 patients with low back pain and 30 pain-free controls. A possible explanation for the high

prevalence of nerve root compression in our MRI reports is that there is no universally accepted imaging criteria to define nerve root compression [27]. When in doubt, radiologists often base their decision regarding the presence of nerve root compression on both the MRI and clinical presentation.
In the present study, the prevalence of serious pathology (3%) might be underestimated, i.e. most patients with suspected serious pathology may have been referred by their GP to hospital, rather than being referred for MRI. On the other hand, the prevalence of serious pathology was three times higher than reported in a study on acute low back patients in pain primary care [9]; this might be due to the differences in patient characteristics between the earlier study (shorter duration of complaints, lower age range) [9] and our study.
The most common serious pathology was vertebral fracture (3%), which is consistent with the findings of Henschke et al. [9]. In our study, malignancy was reported in 5 patients (0.7%) patients, all of whom reported the red flag 'back pain started after age of 50 years' and only 2 patients were known with a history of back pain complaints (40% vs. 81% in the patients without malignancy). A recent review assessed the diagnostic performance of all red flags for malignancy in low back pain [7, 28]. Of the 8 included studies, none included the red flag 'back pain started after age 50 years' despite that this is often seen by GPs as a warning signal.

Our study has several strengths, including a relatively high number of patients and almost no missing data (range 0.3-1%). However, because not all known red flags were included in the questionnaire, we were unable to evaluate the diagnostic rule developed by Henschke et al. [9]. Another limitation is that the presence of several MRI findings might be underestimated (in particular spinal stenosis) due to using the MRI reports instead of standardized scoring of the MR images. However, use of the MRI reports enabled us to reflect daily general practice as closely as possible.

Implications for researchers and clinicians
The GP plays a vital role in early detection of serious diseases. The diagnostic accuracy of red flags for all serious diseases is a challenging topic for future research on low back pain. Evaluation of the performance of different combinations of red flags is recommended, including all relevant red flags currently used in general practice.

Dutch guidelines recommend to refrain from routine immediate lumbar imaging in patients with low back pain and without features suggesting a serious underlying condition [14]. Although a recent review confirmed these recommendations [2], it also stated that these conclusions mainly apply to patients with acute or subacute, non-specific low back pain. Nevertheless, in the present study we observed that the general practitioners in our cohort mostly referred patients with long-lasting severe back complaints, and/or severe sciatica. To date there are no studies on the diagnostic accuracy of MRI in this specific group of patients. Therefore, the use and diagnostic accuracy of MRI in general practice should remain a topic for further research and evaluation.

References

1. Chou, R., R.A. Deyo, and J.G. Jarvik, Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am*, 2012. 50(4): p. 569-85.
2. Chou, R., et al., Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet*, 2009. 373(9662): p. 463-72.
3. van Tulder, M., et al., Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J*, 2006. 15 Suppl 2: p. S169-91.
4. Chou, R., et al., 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*, 2007. 147(7): p. 478-91.
5. Williams, C.M., et al., Low back pain and best practice care: A survey of general practice physicians. *Arch Intern Med*, 2010. 170(3): p. 271-7.
6. Algra, P.R., [MRI diagnostics should be available to general practice] MRI-diagnostiek hoort ook bij de huisarts. *Ned Tijdschr Geneeskd*, 2010. 154: p. A2851.
7. Henschke, N., C.G. Maher, and K.M. Refshauge, Screening for malignancy in low back pain patients: a systematic review. *Eur Spine J*, 2007. 16(10): p. 1673-9.
8. Henschke, N., C.G. Maher, and K.M. Refshauge, A systematic review identifies five "red flags" to screen for vertebral fracture in patients with low back pain. *J Clin Epidemiol*, 2008. 61(2): p. 110-118.
9. Henschke, N., et al., Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. *Arthritis Rheum*, 2009. 60(10): p. 3072-80.
10. Symonds, T.L., et al., Do attitudes and beliefs influence work loss due to low back trouble? *Occup Med (Lond)*, 1996. 46(1): p. 25-32.
11. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy*, 1990. 16(3): p. 199-208.
12. Von Korff, M., M.P. Jensen, and P. Karoly, Assessing global pain severity by self-report in clinical and health services research. *Spine (Phila Pa 1976)*, 2000. 25(24): p. 3140-51.
13. Roland, M. and R. Morris, A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*, 1983. 8(2): p. 141-4.
14. Chavannes AW, M.J., Koes BW, Lubbers WJ, Ostelo R, Spinnewijn WEM, Kolnaar BGM, NHG-Standaard Aspecifieke lagerugpijn. *Huisarts Wet*, 2005. 48(3): p. 113-123.

15. Moons, K.G., et al., Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol*, 2006. 59(10): p. 1092-101.
16. Campbell, P., et al., Prognostic Indicators of Low Back Pain in Primary Care: Five-Year Prospective Study. *J Pain*, 2013.
17. Hill, J.C., et al., Clinical outcomes among low back pain consulters with referred leg pain in primary care. *Spine (Phila Pa 1976)*, 2011. 36(25): p. 2168-75.
18. Mehling, W.E., et al., The prognosis of acute low back pain in primary care in the United States: a 2-year prospective cohort study. *Spine (Phila Pa 1976)*, 2012. 37(8): p. 678-84.
19. Graves, J.M., et al., Early imaging for acute low back pain: one-year health and disability outcomes among Washington State workers. *Spine (Phila Pa 1976)*, 2012. 37(18): p. 1617-27.
20. Jarvik, J.G., et al., Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *JAMA*, 2003. 289(21): p. 2810-8.
21. Jarvik, J.G., et al., Association of Early Imaging for Back Pain With Clinical Outcomes in Older Adults (vol 313, pg 1143, 2015). *Jama-Journal of the American Medical Association*, 2015. 313(17): p. 1758-1758.
22. Jensen, M.C., et al., Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med*, 1994. 331(2): p. 69-73.
23. Kjaer, P., et al., Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine (Phila Pa 1976)*, 2005. 30(10): p. 1173-80.
24. Modic, M.T., et al., Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. *Radiology*, 2005. 237(2): p. 597-604.
25. You, J.J., et al., Patterns of care after magnetic resonance imaging of the spine in primary care. *Spine (Phila Pa 1976)*, 2013. 38(1): p. 51-9.
26. Hancock, M., et al., MRI findings are more common in selected patients with acute low back pain than controls? *Eur Spine J*, 2012. 21(2): p. 240-6.
27. Cihangiroglu, M., et al., Observer variability based on the strength of MR scanners in the assessment of lumbar degenerative disc disease. *Eur J Radiol*, 2004. 51(3): p. 202-8.
28. Henschke, N., et al., Red flags to screen for malignancy in patients with low-back pain. *Cochrane Database Syst Rev*, 2013. 2: p. CD008686.

3

Does MRI add to the prediction of recovery in low back pain patients in general practice?

Evelien I.T. de Schepper, Bart W. Koes, Edwin H.G. Oei, Sita M.A. Bierma-Zeinstra, Pim A.J. Luijsterburg

Abstract

Background

A diverse range of prognostic factors have been studied in relation to low back pain (LBP). Information on the prognostic value of MRI findings could be useful to better inform patients with LBP about their prognosis. However, the prognostic value of MRI findings has mainly been studied in patients with sciatica and these results may differ from studies performed in patients with LBP in general practice.

Objective

To investigate the course and added prognostic value of baseline MRI findings over known prognostic factors for recovery at 12-month follow-up in patients with LBP referred to MRI by their general practitioner.

Methods

Patients (aged ≥ 18 years) referred by their general practitioner for MRI of the lumbar spine were recruited. The questionnaires at baseline and at 3 and 12-months follow-up included potential clinical predictors from history taking and the outcome. The MRI radiology reports were scored for the presence of bulging, disc herniation, nerve root compression, spinal stenosis, spondylolisthesis and serious pathologies. The primary outcome was recovery measured with the Global Perceived Effect scale. Multivariate logistic regression analysis was performed in 3 steps: derivation of a predictive model including characteristics of the patients and back pain only (history taking), including reported MRI findings only, and the addition of reported MRI findings to the characteristics of the patients and back pain.

Results

Pain severity of the patients (n=683) decreased from a mean of 6.6 (SD 2.0) at baseline to 3.8 (SD 2.6) at 3-months follow-up and to 3.8 (SD 2.8) at 12-months follow-up. At 12-months follow-up 53% of the patients reported recovery. Lower age, better attitude/beliefs regarding back pain, acute back pain, presence of neurological symptoms of the leg(s), and presence of non-continuous back pain were significantly associated with recovery at 12-months follow-up: area under the curve (AUC) 0.77. Addition of the MRI findings resulted in an AUC of 0.78.

Conclusions

At 12-months follow-up, only 53% of these patients with low back pain

referred for MRI in general practice reported recovery. Five clinic baseline characteristics were associated with recovery at 12-months follow-up; adding the MRI findings did not result in a stronger prediction of recovery.

Introduction

If a physician suspects the presence of a specific disease, diagnostic imaging can be used. In recent years, general practitioners in the Netherlands can refer low back pain patients for MRI of the lumbar spine themselves. Possible reasons for the use of MRI in general practice are i) to detect or exclude specific pathologies, ii) to reassure the patient (and physician), and/or iii) to prevent unnecessary referrals to secondary care. Despite the recommendations of the guidelines to use MRI only in specific cases, the use of MRI as the initial imaging for low back pain seems to have become more common in general practice in countries such as the USA and Australia [1, 2]. However, data on the use of MRI by general practitioners in the Netherlands are still lacking.

The ultimate goal of any diagnostic test is to improve the clinical outcome of the patient. Well-conducted randomized trials are the top of the diagnostic evidence hierarchy, because they provide the most direct information on the clinical benefits and harms of alternative testing strategies. However, in daily practice most studies on diagnostic tests estimate how accurately they can identify a disease or condition, or how well the test provides prognostic information. Understanding of the prognostic factors in LBP and their relative importance may allow to identify patients who are at a higher risk for developing chronic LBP. Identification of prognostic factors predicting recovery, persistent pain, and disability are important for better understanding of the clinical course, to inform patients and physicians, and to support therapeutic decision making [3]. A diverse range of prognostic factors (demographics, physical factors, and psychological factors) has been studied in relation to persistent LBP [4]. The prognostic value of MRI findings in relation to recovery has mainly been studied in patients with sciatica in secondary care [5-10]; however, these results may differ from studies performed in patients with LBP in general practice.

The aim of this study was to investigate the course and the added prognostic value of baseline MRI findings over known prognostic factors for recovery at 12-months follow-up in patients with LBP referred to MRI by their GP.

Methods

This study is a prospective, observational cohort study in general practice, with a 12-month follow-up. Eligible patients were enrolled between June

2010 and September 2011. The study protocol was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam.

Study population
Consecutive eligible adults referred by their GP for MRI of the lumbar spine were recruited at the MRI Center. The inclusion criteria for the study were: aged ≥ 18 years and referred by their GP for MRI of the lumbar spine. Patients were excluded from the study if there were contraindications for undergoing MRI, or if the patient had insufficient understanding of the Dutch language and/or was incapable of understanding the ramifications of participation.
Eligible patients received written information about the study at the time they made an appointment for their MRI at the MRI Center, and were given the opportunity to ask questions about the study up until the MRI appointment date. When the patient was interested, informed consent was given.

MRI findings
All patients underwent MRI (1.5 Tesla, Siemens, Erlangen, Germany), as scheduled. The MRI protocol consisted of sagittal and transverse T1 and T2 weighted sequences. We performed transverse imaging through affected disks and vertebrae plus a three-dimensional (3D) steady state sequence (CISS). The MRIs were assessed by 1 of 7 radiologists of the MRI Center. As this study was designed to reflect daily general practice as closely as possible, we scored the findings described in the MRI radiology reports retrieved from the MRI Center, which were identical copies of the reports sent to the referring GPs. There was no interference with the care given by the GP or other healthcare providers with respect to advice, diagnostics or treatment.
A single reader [EdS], who was trained by a radiologist [EO] and blinded to the participants' clinical data, extracted data from the MRI reports regarding the presence or absence of the following findings at each lumbar level (T12-L1 through L5-S1): intervertebral disc bulging, disc herniation (protrusion/extrusion), nerve root compression, spinal stenosis, spondylolisthesis, and serious pathology (fracture, malignancy and/or discitis).

Outcomes and potential predictors
After inclusion, the baseline measurement included validated questionnaires, for which participants were invited by email containing a secured

link to the online questionnaires. The follow-up period was 12 months, with follow-up measurements at 3 and 12 months. Reminders were sent by email after 2 and 3 weeks of non-response.
The primary outcome measure was recovery, defined as a score of 'strongly improved' or 'completely recovered' on the Global Perceived Effect (GPE) scale [11]. No recovery was defined as a GPE score of 'somewhat improved', 'stayed the same', 'somewhat worsened', 'strongly worsened', or 'worse than ever'. Secondary outcome measures included severity of back pain measured on an 11-point numerical rating scale (NRS) in which 0 represents 'no pain' and 10 represents 'unbearable pain' [12]; and disability was measured using the Roland Disability Questionnaire (RDQ), with scores ranging from 0 (no disability) to 24 (severe disability) [13]. Recovery of the secondary outcomes was defined as 'severity of back pain <3 (NRS)' or 'disability score <4 (RDQ)' at 12-months follow-up [14].

The baseline questionnaire included measurements of potential predictors for recovery. We chose 21 candidate predictors reported to be prognostic and/or deemed clinically relevant, taking into account the rule of thumb that logistic regression models require a minimum of 10 events per predictor [15]. These factors were divided into three categories: 1) Patient characteristics: age, gender, body mass index (BMI), level of education, employment status, and attitude/beliefs about low back pain at baseline (BBQ, range 9-45) [16]; 2) Back pain characteristics: duration of back symptoms, history of back pain, severity of back pain at baseline (NRS), presence of radiating pain in the legs below the knee, neurological symptoms of the legs, morning stiffness of the back, presence of continuous back pain independent of posture or activity, disability at baseline (RDQ, range 0-24), and history of back surgery; 3) MRI findings: bulging, disc herniation, nerve root compression, spinal stenosis, spondylolisthesis, and serious pathology (fracture, malignancy and/or discitis). Neurological symptoms were determined with the question: "*Did you have any complaints of numbness or tingling of the leg(s), and/or weakness of the leg(s) during the last week*" (answer yes/no).

Statistical analysis
Descriptive analyses were used to report the characteristics of the patients and the course of back pain over the 12-month follow-up period using the mean and standard deviation (SD) for continuous data, and proportions for categorical data. Data were screened for inconsistencies and missing

baseline data were imputed using multivariate imputation resulting in 5 imputed datasets [17]. Visual inspection of the linear relationship of all continuous variables and the primary outcome revealed nonlinearity between BMI and RDQ score at baseline with the outcome. Therefore, BMI was dichotomized into <25 and ≥25 and RDQ score into <18 and ≥18. To enable easy interpretation of predictors in a clinical setting, we dichotomized the following categorical variables: education was dichotomized in low (lower secondary school or compulsory education) and high level education; and duration of back pain in acute (≤ 3 months) and chronic back pain.

A correlation matrix was observed for all potential predictors to check for co-linearity, setting the cut-off value for Pearson's correlation coefficient (R) at 0.70. None of the predictors were highly correlated. Multiple (backward) logistic regression analyses were performed (entry 0.05, removal 0.10) to determine which baseline factors were associated with the primary outcome.

The analyses were carried out in 3 steps: derivation of a predictive model 1) including patients' and back pain characteristics only (history taking), 2) including reported MRI findings only, 3) including both patients' and back pain characteristics and MRI findings. If potential predictors were selected in at least 3 of 5 imputed databases in the multivariate analysis, they were included in the final model (enter method; $p < 0.05$). To evaluate the discriminative ability of the models, a receiver operating characteristic curve was generated for the predicted probabilities and the area under the curve (AUC) was calculated [18]. The predictive value of the MRI findings was evaluated by observing the increase in discriminative ability (AUC) with the DeLong test [19].

Additional exploratory analyses were carried out: 1) in the subgroup of patients with indications for which clinical practice guidelines recommend imaging (specified as patients with pain radiating in the leg below the knee (≥7 NRS) for ≥ 6 weeks at baseline), 2) in the subgroup of patients without surgery during the 12-month follow-up, and 3) with the secondary outcomes 'severity of back and leg pain <3 (NRS)' and 'disability score <4 (RDQ)' at 12-months follow-up. Analyses were performed using SPSS version 20.0 (SPSS Inc., USA) and MedCalc version 12.4.0.0 (MedCalc Software bvba).

Results

A total of 683 referred patients participated in the study (Figure 1). During follow-up, 547 (80%) patients returned the 3-month follow-up questionnaire and 474 (69%) returned the 12-month follow-up questionnaire. Information on BMI at baseline was missing in 8 patients (1%).
The baseline characteristics of the patients are presented in Table 1. The mean age of the patients was 49.9 (SD 12.5; range 19-80) years. In total, 53% of the patients were male. At baseline, 33% of the patients reported acute back pain. Of all patients, 66% reported radiating pain in the leg below the knee; 77% reported neurological symptoms of the leg(s).
The MRI reports described disc herniation in 72% of the patients; 69% of the MRI reports mentioned signs of nerve root compression. Spinal stenosis was reported in 13% of the patients. Serious pathologies (fractures, malignancies and discitis) were reported in 3% of the patients.

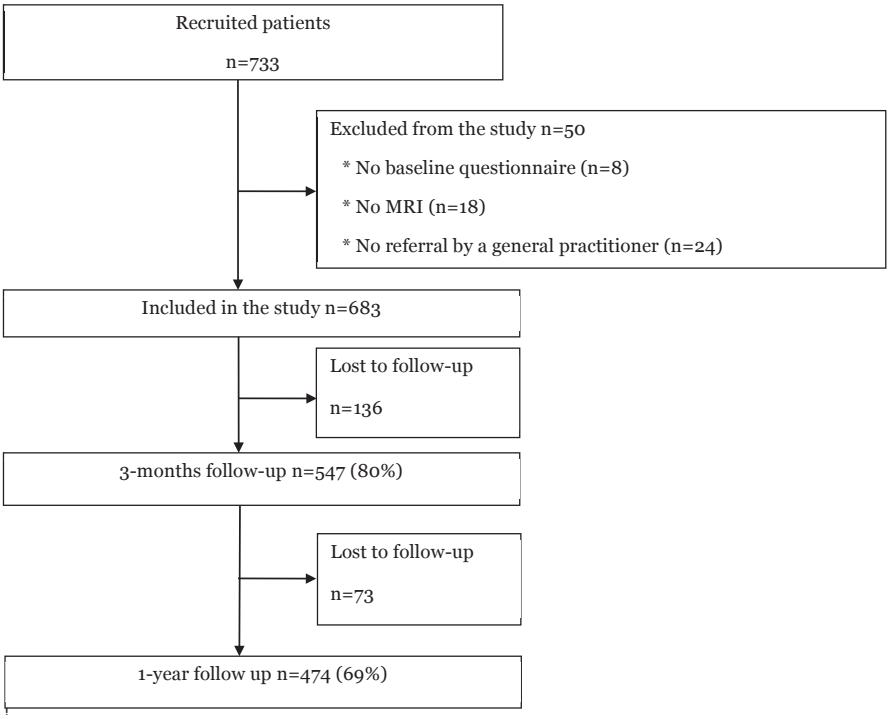

Figure 1. Flow chart of the study population

Table 1 Baseline characteristics of the included 683 patients and of the 251 patients that reported recovery at 12 months follow-up*

	Study population *n*=683		**Recovered at 12 months *n*=251**	
Patient characteristics				
Age in years, mean (SD)	49.9	(12.5)	49.8	(12.4)
Male	365	(53)	146	(58)
BMI, mean (SD)	25.9	(3.8)	26.0	(3.6)
BMI ≥25	390	(57)	150	(60)
Education level low	244	(36)	72	(29)
Employed (paid job)	479	(70)	183	(73)
Attitude and beliefs about back pain (BBQ), mean (SD)	26.3	(6.1)	28.0	(5.8)
Back pain characteristics				
Acute back pain (<3 months)	228	(33)	118	(47)
History of back pain	549	(80)	198	(79)
Severity of back pain (NRS), mean (SD)	6.6	(2.0)	6.1	(2.3)
Pain radiating in the leg below the knee	450	(66)	185	(74)
Neurological symptoms in legs	525	(77)	206	(82)
Morning stiffness of the back	353	(52)	117	(47)
Continuous back pain	347	(51)	91	(36)
Disability (RDQ), mean (SD)	13.5	(5.2)	13.6	(4.8)
RDQ ≥18	173	(25)	58	(23)
History of back surgery	112	(16)	38	(15)
MRI findings				
Bulging	308	(45)	109	(43)
Disc herniation	492	(72)	200	(80)
Nerve root compression	472	(69)	200	(80)
Spinal stenosis	87	(13)	29	(12)
Spondylolisthesis	56	(8)	18	(7)
Serious pathology**	22	(3)	9	(4)

*Data are presented as numbers (percentages) unless otherwise indicated
**Serious pathology: impression fracture, malignancy and/or discitis
SD: standard deviation; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/belief regarding back pain;
NRS: numeric rating scale (range 0-10, 0 means no pain); RDQ: Roland disability questionnaire (range 0-24), a higher score indicates worse health

Table 2 Outcomes at baseline, and at 3 and 12-months follow-up

	Baseline (*n*=683)	**3-months follow-up (*n*=547)**	**12-months follow-up (*n*=474)**
Recovery (GPE), *n* (%)	-	240 (44)	251 (53)
Severity of back pain (NRS), mean (SD)	6.6 (2.0)	3.8 (2.6)	3.8 (2.8)
Disability (RDQ), mean (SD)	13.5 (5.2)	8.5 (6.0)	6.5 (5.7)

GPE: Global Perceived Effect; 7-point Likert scale, dichotomized in 1-2: recovery, 3-7: no recovery; NRS: numeric rating scale (range 0-10, 0 means no pain); RDQ: Roland disability questionnaire (range 0-24, 0 means no disability), a higher score indicates worse health; SD: standard deviation

Course

At 3-months follow-up, the mean back pain severity had decreased from 6.6 (SD 2.0) to 3.8 (SD 2.6) on the 11-point NRS, and 44% of the patients reported recovery. At 12-months follow-up, the mean back pain severity was 3.8 (SD 2.8), and 53% of the patients reported recovery. The mean disability score was 13.5 (SD 5.2) at baseline, 8.5 (SD 6.0) at 3-months, and 6.5 (SD 5.7) at 12-months follow-up (Table 2).

Predictors of recovery

Table 3 shows the results of the multivariate logistic regression analysis regarding the potential predictors on the primary outcome recovery.
In the first model that included patient and back pain characteristics as potential predictors, the variables associated with recovery were: age [odds ratio (OR) 0.98; 95% confidence interval (CI): 0.96-0.99], the BBQ score (OR 1.1; CI: 1.0-1.1), acute back pain (OR 3.0; CI: 1.9-4.8), neurological symptoms of the leg(s) (OR 2.3; CI: 1.4-3.9), and continuous back pain independent of posture or activity (OR 0.3; CI: 0.2-0.5). The AUC for this model was 0.77 (Table 3).
The second model was calculated with the MRI findings. The variables associated with recovery were: discus hernia (OR 1.6; CI 1.1-2.6), and nerve root compression (OR 2.2; CI 1.4-3.5). The AUC for this model was 0.63.
When model 2 (the MRI findings) was added to the first model, the AUC increased to 0.78 and the variables associated with recovery were: age (OR 0.98; CI: 0.96-1.0), the BBQ score (OR 1.1; CI: 1.0-1.1), acute back pain (OR 2.8; CI: 1.7-4.5), neurological symptoms of the legs (OR 2.0; CI: 1.2-3.5), continuous back pain independent of posture or activity (OR 0.3; CI: 0.2-0.5), and nerve root compression (OR 2.2; CI: 1.4-3.4). The discriminative ability (AUC) of model 1 and 3 showed no significant difference (p=0.086).

Table 3 Results of multivariate logistic regression analysis regarding potential predictors and recovery at 12-months follow-up (*n*=474)

Recovery (GPE)	Pooled OR (95% CI)	p-value	AUC
1. Patient and back pain characteristics			**0.77**
Age	0.98 (0.96 - 0.99)	< 0.01	
Attitude and beliefs about back pain (BBQ)	1.1 (1.0 - 1.1)	< 0.01	
Acute pain at baseline (yes)	3.0 (1.9 - 4.8)	< 0.01	
Neurological symptoms in legs (yes)	2.3 (1.4 - 3.9)	< 0.01	
Continuous back pain (yes)	0.3 (0.2 - 0.5)	< 0.01	
2. MRI findings			
Disc herniation (yes)	1.6 (1.1 - 2.6)	< 0.05	**0.63**
Nerve root compression (yes)	2.2 (1.4 - 3.5)	< 0.01	
3. Patient and back pain characteristics + MRI findings			
Age	0.98 (0.96 - 1.0)	< 0.05	
Attitude and beliefs about back pain (BBQ)	1.1 (1.0 - 1.1)	< 0.01	**0.78**
Acute pain at baseline (yes)	2.8 (1.7 - 4.5)	< 0.01	
Neurological symptoms of legs (yes)	2.0 (1.2 - 3.5)	< 0.01	
Continuous tback pain (yes)	0.3 (0.2 - 0.5)	< 0.01	
Nerve root compression (yes)	2.2 (1.4 - 3.4)	< 0.01	

GPE: Global Perceived Effect; 7-point Likert scale, dichotomized in 1-2: recovery, 3-7: no recovery OR: odds ratio; AUC: area under the curve; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; MRI: magnetic resonance imaging

Additional exploratory analyses

One of the main groups for which clinical practice guidelines recommend imaging is the group of sciatica patients with an indication for surgery, specified as sciatica patients with severe pain for ≥ 6 weeks. Additional analyses in this group of patients (n=259) showed an AUC of 0.78 for the first model (Supplemental Digital Content 1). When the MRI findings were added to the first model, the AUC remained 0.78. Again, the discriminative ability (AUC) of model 1 and model 3 showed no significant difference (p= >0.05).

To study the possible influence of surgery during follow-up, additional analyses in patients without surgery during follow-up (n=559) were performed (Supplemental Digital Content 2). The analyses showed an AUC of 0.80 for the first model. When the MRI findings were added to the first model, the AUC (0.80) showed no significant difference.

Secondary analyses with the outcome 'severity of back and leg pain <3 (NRS)' at 12-months follow-up showed an AUC of 0.73 for the first model

(Supplemental Digital Content 3). When the MRI findings were added to the first model, the AUC (0.74) showed no significant difference. Secondary analyses with the outcome 'disability score < 4 (RDQ)' at 12-months follow-up showed an AUC of 0.76 for the first model (Supplemental Digital Content 4). When the variables of the MRI findings were added (model 3), none of the MRI findings were significant predictors.

Discussion

This study presents the course of low back pain in 683 patients who were referred for MRI of the lumbar spine by their GP and identified predictors for recovery at 12-months follow-up. Back pain severity of the patients decreased from a mean of 6.6 (SD 2.0) at baseline to 3.8 (SD 2.6) at 3-months follow-up and to 3.8 (SD 2.8) at 12-months follow-up. At 12-months follow-up 53% of the patients reported recovery. Lower age, better attitude/beliefs regarding back pain, acute back pain, presence of neurological symptoms of the leg(s), and presence of non-continuous back pain were significantly associated with recovery at 12-months follow-up (AUC 0.77). Addition of the reported MRI findings did not add to the predictive value of the prognostic model with clinical factors only.

To our knowledge, this is the first prospective cohort study of patients with LBP referred for MRI in a primary care setting. Baseline back pain severity scores were higher than reported in earlier LBP cohort studies [4, 20-22]. Disability scores are similar to those in two studies that included patients with LBP referred for MRI or radiography by their primary physician [20, 23].
Back pain severity mainly decreased during the first 3 months and then remained relatively stable between 3 and 12 months. A similar pattern was found in other (back) pain studies [23-26]. In the review by Pengel et al. only studies investigating patients with acute back pain were included [24]. In our cohort study the pattern was also visible in patients reporting chronic back pain. A possible explanation for this observation could be that the chronic back pain patients visited their GP during a flare-up of their back pain and therefore showed a pain pattern similar to patients with acute back pain. Another explanation for the improvement of patients over time may be regression to the mean, which is a consequence of random variation over time [26].

In our cohort, MRI reports showed disc herniation in 72% and nerve root compression in 69% of the patients. Both these prevalences are higher than reported in other cohorts that included patients with LBP [23, 27]. As expected, of the serious pathologies (fractures, malignancies and discitis), the most frequently observed serious pathology was vertebral fracture (3%). This is consistent with a recent study on the prevalence of serious spinal pathology in primary care [28].

In the field of LBP, previous studies presented inconsistent conclusions regarding important prognostic factors for recovery [29]. Only a small number of important prognostic factors were consistently reported; of these, both lower age and acute back pain were also related to recovery at follow-up in our cohort. Negative beliefs about LBP was only reported in one other study as an independent risk factor for poor recovery [30], and was associated with high back pain intensity levels in a cross-sectional study [31]. Continuous back pain was reported as a factor for poor recovery in only one study [32]. The question about the presence of continuous back pain independent of posture or activity is often used in primary care, but is not often examined in prognostic studies.

A recent review reported that the presence of pain radiating down the leg, with neurological findings, was associated with a poor prognosis in patients with LBP [33]. In our model, neurological symptoms of the legs were positively associated with recovery. An explanation for this could be that the included patients without neurological symptoms of the legs tend to be worse off in terms of pain, disability and duration of complaints.
The AUC of the multiple regression model remained similar when the variables of the MRI findings were added to the model that included characteristics of the patients and of back pain. This indicates no additional value of the reported MRI findings with regard to the discriminative value to predict recovery at 12-months follow-up. The only MRI finding that remained in the model was 'nerve root compression' and, when it is was included, the association of the variables 'acute pain' and 'neurological symptoms' diminished.

Strengths of the present study are that it included a relatively high number of patients, had low dropout rates despite the use of online questionnaires, and had almost no missing data (1%). However, some limitations need to be considered when interpreting the results. One limitation is the pos-

sibility of selection bias due to the large number of patients (n=2242) not willing to participate. However, the characteristics of this latter group (e.g. gender, age, living district) showed no significant difference compared with those of the included patients.
Also, the presence of several MRI findings might be underestimated (in particular spinal stenosis) due to using the MRI reports instead of standardized scoring of the MR images. Further research may be needed to assess the discriminative value of systematically scored MRIs. However, use of the MRI reports reflects daily general practice as closely as possible.

Implications for clinicians
Understanding of the prognostic factors in LBP and their relative importance may allow to identify patients at a higher risk for developing chronic complaints. Predictors for recovery were lower age, acute back pain at baseline, the presence of neurological symptoms of the leg(s), the presence of non-continuous back pain, and better attitude/beliefs regarding back pain. Adding MRI findings did not result in a stronger prediction of recovery at 12-months follow-up.
These findings suggest that GPs can provide a moderately good prediction of the prognosis of their patients with LBP based on their characteristics and complaints (history taking); information from the MRI reports does not offer added prognostic value. Although MRI reports do not provide additional prognostic value regarding recovery, the real diagnostic accuracy of lumbar MRI in this group of patients is still unknown.

References

1. Chou, R., R.A. Deyo, and J.G. Jarvik, Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am*, 2012. 50(4): p. 569-85.
2. Williams, C.M., et al., Low Back Pain and Best Practice Care A Survey of General Practice Physicians. *Archives of Internal Medicine*, 2010. 170(3): p. 271-277.
3. Verwoerd, A.J., et al., Systematic review of prognostic factors predicting outcome in non-surgically treated patients with sciatica. *Eur J Pain*, 2013. 17(8): p. 1126-37.
4. Campbell, P., et al., Prognostic Indicators of Low Back Pain in Primary Care: Five-Year Prospective Study. *J Pain*, 2013.
5. Carragee, E.J. and D.H. Kim, A prospective analysis of magnetic resonance imaging findings in patients with sciatica and lumbar disc herniation. Correlation of outcomes with disc fragment and canal morphology. *Spine (Phila Pa 1976)*, 1997. 22(14): p. 1650-60.
6. Jensen, T.S., et al., Magnetic resonance imaging findings as predictors of clinical outcome in patients with sciatica receiving active conservative treatment. *J Manipulative Physiol Ther*, 2007. 30(2): p. 98-108.
7. Komori, H., et al., Factors predicting the prognosis of lumbar radiculopathy due to disc herniation. *J Orthop Sci*, 2002. 7(1): p. 56-61.
8. Peul, W.C., et al., Influence of gender and other prognostic factors on outcome of sciatica. *Pain*, 2008. 138(1): p. 180-91.
9. Takada, E., M. Takahashi, and K. Shimada, Natural history of lumbar disc hernia with radicular leg pain: Spontaneous MRI changes of the herniated mass and correlation with clinical outcome. *J Orthop Surg (Hong Kong)*, 2001. 9(1): p. 1-7.
10. Vroomen, P.C., J.T. Wilmink, and K.M. de, Prognostic value of MRI findings in sciatica. *Neuroradiology*, 2002. 44(1): p. 59-63.
11. Beurskens, A.J., H.C. de Vet, and A.J. Koke, Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*, 1996. 65(1): p. 71-6.
12. Von Korff, M., M.P. Jensen, and P. Karoly, Assessing global pain severity by self-report in clinical and health services research. *Spine (Phila Pa 1976)*, 2000. 25(24): p. 3140-51.
13. Roland, M. and R. Morris, A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*, 1983. 8(2): p. 141-4.

14. Mehling, W.E., et al., Acute low back pain and primary care: how to define recovery and chronification? *Spine (Phila Pa 1976)*, 2011. 36(26): p. 2316-23.
15. Peduzzi, P., et al., A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*, 1996. 49(12): p. 1373-9.
16. Symonds, T.L., et al., Do attitudes and beliefs influence work loss due to low back trouble? *Occup Med (Lond)*, 1996. 46(1): p. 25-32.
17. Moons, K.G., et al., Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol*, 2006. 59(10): p. 1092-101.
18. Harrell, F.E., Jr., Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. 2001, New York: Springer.
19. DeLong, E.R., D.M. DeLong, and D.L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 1988. 44(3): p. 837-45.
20. Graves, J.M., et al., Early imaging for acute low back pain: one-year health and disability outcomes among Washington State workers. *Spine (Phila Pa 1976)*, 2012. 37(18): p. 1617-27.
21. Hill, J.C., et al., Clinical outcomes among low back pain consulters with referred leg pain in primary care. *Spine (Phila Pa 1976)*, 2011. 36(25): p. 2168-75.
22. Mehling, W.E., et al., The prognosis of acute low back pain in primary care in the United States: a 2-year prospective cohort study. *Spine (Phila Pa 1976)*, 2012. 37(8): p. 678-84.
23. Jarvik, J.G., et al., Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *JAMA*, 2003. 289(21): p. 2810-8.
24. Pengel, L.H., et al., Acute low back pain: systematic review of its prognosis. *BMJ*, 2003. 327(7410): p. 323.
25. Suri, P., et al., Acute low back pain is marked by variability: An internet-based pilot study. *BMC Musculoskelet Disord*, 2011. 12: p. 220.
26. da, C.M.C.L., et al., The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ*, 2012. 184(11): p. E613-24.
27. Modic, M.T., et al., Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. *Radiology*, 2005. 237(2): p. 597-604.
28. Henschke, N., et al., Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. *Arthritis Rheum*, 2009. 60(10): p. 3072-80.

29. Hayden, J.A., et al., Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*, 2009. 62(8): p. 781-796 e1.
30. Elfering, A., et al., Beliefs about back pain predict the recovery rate over 52 consecutive weeks. *Scand J Work Environ Health*, 2009. 35(6): p. 437-45.
31. Urquhart, D.M., et al., Negative beliefs about low back pain are associated with high pain intensity and high level disability in community-based women. *BMC Musculoskelet Disord*, 2008. 9: p. 148.
32. Harms, M.C., C.E. Peers, and D. Chase, Low back pain: what determines functional outcome at six months? An observational study. *BMC Musculoskelet Disord*, 2010. 11: p. 236.
33. Konstantinou, K., et al., The impact of low back-related leg pain on outcomes as compared with low back pain alone: a systematic review of the literature. *Clin J Pain*, 2013. 29(7): p. 644-54.

Supplemental Digital Content 1 Results of multivariate logistic regression analysis regarding potential predictors and recovery at 12-months follow-up in sciatica patients with an indication for surgery* (*n*=176)

Recovery (GPE)	Pooled OR (95% CI)		p-value	AUC
1. Patient and back pain characteristics				**0.78**
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.2)	< 0.01	
Acute pain at baseline (yes)	4.4	(1.7 - 11.0)	< 0.01	
Severity of back pain (NRS)	0.6	(0.4 - 0.9)	< 0.05	
Morning stiffness of the back (yes)	2.1	(1.0 - 4.4)	< 0.05	
Continuous back pain (yes)	0.5	(0.2 - 1.0)	0.05	
2. MRI findings				**0.60**
Nerve root compression (yes)	3.0	(1.4 - 6.4)	< 0.01	
3. Patient and back pain characteristics + MRI findings				**0.78**
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.2)	< 0.01	
Acute pain at baseline (yes)	4.8	(2.0 - 11.4)	< 0.01	
Severity of back pain (NRS)	0.5	(0.3 - 0.8)	< 0.01	
Nerve root compression	3.1	(1.3 - 7.3)	< 0.01	

*Indication for surgery: patients with pain radiating in the leg below the knee (≥7 NRS) for more than 6 weeks

GPE: Global Perceived Effect; 7-point Likert scale, dichotomized in 1-2: recovery, 3-7: no recovery; OR: odds ratio; AUC: area under the curve; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; NRS: numeric rating scale (range 0-10, 0 means no pain); MRI: magnetic resonance imaging

Supplemental Digital Content 2 Results of multivariate logistic regression analysis regarding potential predictors and recovery at 12-months follow-up in patients without surgery during follow-up (*n*=366)

Recovery (GPE)	Pooled OR (95% CI)		p-value	AUC
1. Patient and back pain characteristics				**0.80**
Age	0.97	(0.95 - 1.0)	< 0.05	
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.1)	< 0.01	
Acute pain at baseline (yes)	3.9	(2.3 - 6.6)	< 0.01	
Severity of back pain (NRS)	0.9	(0.8 - 1.0)	< 0.05	
Neurological symptoms in legs	1.9	(1.1 - 3.4)	< 0.05	
Continuous back pain (yes)	0.4	(0.2 - 0.6)	< 0.01	
2. MRI findings				**0.63**
Disc herniation (yes)	1.9	(1.1 - 3.1)	< 0.05	
Nerve root compression (yes)	1.9	(1.2 - 3.1)	< 0.01	
3. Patient and back pain characteristics + MRI findings				**0.80**
Age	0.98	(0.96 - 1.0)	< 0.05	
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.1)	< 0.01	
Acute pain at baseline (yes)	3.5	(2.1 - 5.9)	< 0.01	
Continuous back pain (yes)	0.3	(0.2 - 0.5)	< 0.01	
Nerve root compression (yes)	2.1	(1.3 - 3.6)	< 0.01	
Spinal stenosis (yes)	0.4	(0.1 - 0.9)	< 0.05	

GPE: Global Perceived Effect; 7-point Likert scale, dichotomized in 1-2: recovery, 3-7: no recovery OR: odds ratio; AUC: area under the curve; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; NRS: numeric rating scale (range 0-10, 0 means no pain); MRI: magnetic resonance imaging

Supplemental Digital Content 3 Results of multivariate logistic regression analysis regarding potential predictors and the secondary outcome 'severity of back and leg pain <3 (NRS)' at 12-months follow-up (*n*=454)

Severity back and leg pain <3 (NRS)	Pooled OR (95% CI)		p-value	AUC
1. Patient and back pain characteristics				**0.73**
Gender	2.0	(1.3 - 3.0)	< 0.01	
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.1)	< 0.05	
Acute pain at baseline (yes)	2.1	(1.4 - 3.3)	< 0.01	
Pain radiating in the leg below the knee (yes)	1.7	(1.1 - 2.8)	< 0.05	
Continuous back pain (yes)	0.5	(0.3 - 0.8)	< 0.01	
2. MRI findings				**0.58**
Disc herniation (yes)	2.5	(1.5 - 4.0)	< 0.01	
3. Patient and back pain characteristics + MRI findings				**0.74**
Gender	1.8	(1.2 - 2.8)	< 0.01	
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.1)	< 0.05	
Acute pain at baseline (yes)	2.0	(1.3 - 3.2)	< 0.01	
Pain radiating in the leg below the knee (yes)	1.7	(1.1 - 2.7)	< 0.05	
Continuous back pain (yes)	0.5	(0.3 - 0.8)	< 0.01	
Disc herniation (yes)	1.8	(1.1 - 3.0)	< 0.05	

NRS: numeric rating scale (range 0-10, 0 means no pain); OR: odds ratio; AUC: area under the curve; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; MRI: magnetic resonance imaging

Supplemental Digital Content 4 Results of multivariate logistic regression analysis regarding potential predictors and the secondary outcome 'disability score <4 (RDQ)' at 12-months follow-up (*n*=474)

Severity back pain <3 (NRS)	Pooled OR (95% CI)		p-value	AUC
1. Patient and back pain characteristics				**0.76**
Employed (yes)	1.9	(1.2 - 3.0)	< 0.01	
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.1)	< 0.01	
Acute pain at baseline (yes)	2.8	(1.7 - 4.4)	< 0.01	
Continuous back pain (yes)	0.5	(0.3 - 0.7)	< 0.01	
Disability ≥18 (RDQ)	0.5	(0.3 - 0.9)	< 0.05	
History of back surgery (yes)	0.5	(0.3 - 0.8)	0.01	
2. MRI findings				
-	-		-	
3. Patient and back pain characteristics + MRI findings				**0.76**
Employed (yes)	1.9	(1.2 - 3.0)	< 0.01	
Attitude and beliefs about back pain (BBQ)	1.1	(1.0 - 1.1)	< 0.01	
Acute pain at baseline (yes)	2.8	(1.7 - 4.4)	< 0.01	
Continuous back pain (yes)	0.5	(0.3 - 0.7)	< 0.01	
Disability ≥18 (RDQ)	0.5	(0.3 - 0.9)	< 0.05	
History of back surgery (yes)	0.5	(0.3 - 0.8)	0.01	

RDS: Roland disability questionnaire (range 0-24), a higher score indicates worse health; NRS: numeric rating scale (range 0-10, 0 means no pain); OR: odds ratio; AUC: area under the curve; BBQ: Back beliefs questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; MRI: magnetic resonance imaging

4

Healthcare service use after lumbar spine MRI in general practice

Evelien I.T. de Schepper, Bart W. Koes, Erik F.H. Veldhuizen, Edwin H.G. Oei, Sita M.A. Bierma-Zeinstra, Pim A.J. Luijsterburg

Abstract

Design

Observational prospective cohort study with a 12-month follow-up.

Background

An understanding of the patterns of healthcare services used after MRI of the spine in general practice would provide information about how MRI scans are used in primary care. Identification of possible prognostic factors predicting healthcare use can be important to inform patients and physicians.

Objective

To investigate the association between patient characteristics, back pain characteristics and MRI abnormalities with subsequent specialist consultation and/or surgery in low back pain patients referred to MRI by their general practitioner.

Methods

Patients (aged 18 years and over) referred by their general practitioner for MRI of the lumbar spine were recruited. The MRI radiology reports were scored regarding the presence of bulging, disc herniation, nerve root compression, spinal stenosis, spondylolisthesis and serious pathologies (i.e. fracture, malignancy, discitis). The questionnaires filled in at baseline, and at 3 and 12-month follow-up, included potential clinical predictors from history taking and use of healthcare services.

Results

Of the 683 included patients, 301 (55%) reported consultation with a specialist during the first 3 months, and 124 (18%) underwent spine surgery during the 12-month follow-up. Five clinical baseline characteristics were associated with consultation, including four characteristics from history taking. Younger patients, with pain radiating in the leg below the knee, severe disability, a history of back surgery, presence of nerve root compression or spinal stenosis on MRI were more likely to undergo subsequent spine surgery (AUC 0.75).

Conclusions

At 12-month follow-up, 18% of patients with low back pain referred for MRI in general practice underwent surgery. Six baseline characteristics were associated with surgery during follow-up; including nerve root compression and spinal stenosis on MRI.

Introduction

Dutch general practitioners (GPs) can directly refer low back pain patients for magnetic resonance imaging (MRI) of the lumbar spine, the use of which has become increasingly routine [1, 2]. MRI can provide anatomical information to guide subsequent management when it is used in the appropriate clinical context. Clinical practice guidelines recommend to reserve MRI for patients with a suspected serious underlying condition or neurological deficits or who are candidates for surgery [1, 3-6]. Other possible reasons for the use of MRI in general practice are to reassure the patient (and physician), or to prevent unnecessary referrals to secondary care.

Data on the patterns of subsequent care among patients referred for lumbar spine MRI by their GP are scarce. Recently, You et al. [7] reported the results of a Canadian retrospective study examining use of healthcare services after MRI of the spine requested by a primary care physician. The authors determined that MRI scans of the spine performed in symptomatic patients in primary care showed a high prevalence of abnormalities, and that the results of the MRI scans were not strongly predictive of subsequent surgery.

A better understanding of the characteristics of the patients and of the patterns of healthcare services used after MRI of the lumbar spine would provide information about how MRI findings are used by GPs for subsequent management. Identification of possible prognostic factors predicting consultation with specialists or surgery can be important to inform patient and physician.

This study aimed to investigate the association between specific patient characteristics, back pain characteristics and MRI abnormalities with subsequent specialist consultation during 3-month follow-up and/or surgery during 12-month follow-up in low back pain patients referred to MRI by their GP.

Methods

This study is a prospective, observational cohort study in general practice, with a 12-month follow-up. Eligible patients were enrolled between June 2010 and September 2011. The study protocol was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, the Netherlands.

Study Population

Consecutive eligible adults referred by their GP for MRI of the lumbar spine were recruited at the MRI Center, the Netherlands. The inclusion criteria for the cohort study were: an age of 18 years or over and being referred by their GP for MRI of the lumbar spine. Patients were excluded from the study if there were contra-indications for undergoing MRI, or if the patient had insufficient understanding of the Dutch language and/or was incapable of understanding the ramifications of participation.
Eligible patients received written information about the study at the time they made an appointment for their MRI at the MRI Center, and were given the opportunity to ask questions about the study up until the MRI appointment date. When the patient was interested, informed consent was provided by the patient.

MRI findings

All patients underwent MRI (1.5 Tesla, Siemens, Erlangen, Germany), as scheduled. The MRI protocol consisted of sagittal and transverse T1 and T2-weighted sequences. We performed transverse imaging through affected disks and vertebrae plus a three-dimensional (3D) steady state sequence (CISS). The MRIs were assessed by one of 7 radiologists of the MRI Center. As this study was designed to reflect daily general practice as closely as possible, we scored the findings described in the MRI radiology reports retrieved from the MRI Center, which were identical copies of the reports sent to the referring GP. There was no interference with the care given by the GP or other healthcare providers with respect to advice, diagnostics or treatment.
A single reader [EdS], who was trained by a radiologist [EO] and blinded to the participants' clinical data, extracted data from the MRI reports regarding the presence or absence of the following findings at each lumbar level (T12-L1 through L5-S1): intervertebral disc bulging, disc herniation (protrusion/extrusion), disc sequestration, nerve root compression, spinal stenosis, spondylolisthesis, fracture, malignancy and discitis.

Outcomes and potential predictors

After inclusion, the baseline measurement included validated questionnaires, for which participants were invited by email containing a secured link to the online questionnaire. The follow-up period was 12 months, with follow-up measurements at 3 and at 12 months. Reminders were sent by email after 2 and 3 weeks of non-response.

The baseline questionnaire included measurements of potential predictors for subsequent healthcare use. We selected 24 candidate predictors that had previously been reported to be prognostic and/or deemed clinically relevant [7-12]. These factors were divided into three categories: 1) Patient characteristics: age, gender, body mass index, level of education, employment status, and attitude and beliefs about low back pain measured with the Back Beliefs Questionnaire (BBQ, range 9-45) [13] in which a higher score indicates better attitude/beliefs regarding back pain; 2) Back pain characteristics: history of back pain, duration of back symptoms, severity of back pain measured on an 11-point numerical rating scale (NRS) in which 0 represents 'no pain' and 10 represents 'the worst pain ever', presence of radiating pain in the legs below the knee, severity of leg pain (NRS), neurological symptoms of the legs, morning stiffness of the back, presence of continuous back pain independent of posture or activity, disability measured with the Roland Disability Questionnaire (RDQ, range 0-24), history of back surgery, recent back surgery and recent consultation with a specialist; 3) MRI findings: intervertebral disc bulging, disc herniation, nerve root compression, spinal stenosis, spondylolisthesis, and serious pathology (fracture, malignancy and/or discitis).

Neurological symptoms were determined with the question: "Did you have any complaints of numbness or tingling of the leg(s), and/or weakness of the leg(s) during the last week" (answer yes/no). Recent back surgery was defined as back surgery in the year prior to the baseline measurement. Recent consultation with a specialist was defined as a consultation with an orthopedic surgeon, neurologist, neurosurgeon or rheumatologist in the 3 months prior to the baseline measurement.
The follow-up questionnaires at 3 and 12 months included the following types of healthcare use due to back pain: 1) medication used in the previous 3 months, 2) consultation with a GP, physical therapist, specialist (orthopaedic surgeon, neurologist, neurosurgeon, rheumatologist), occupational physician, psychologist or multidisciplinary pain team in the previous 3 months, and 3) surgery during follow-up. In the Netherlands, patients cannot consult a specialist without a referral from their general practitioner.

Statistical analyses
Descriptive analyses were used to report the characteristics of the patients and the healthcare use at baseline, at 3-months and at 12-months follow-up using the mean and the SD for continuous data, and proportions for

categorical data. Data were screened for inconsistencies and missing baseline data were imputed using multivariate imputation resulting in 5 imputed datasets [14]. To enable easy interpretation of predictors in a clinical setting, we dichotomized the following categorical variables: education was dichotomized in low (lower secondary school or compulsory education) and high level education; and duration of back pain in acute (less than 3 months) and chronic back pain.
A correlation matrix was examined for all potential predictors to check for co linearity, setting the cutoff value for Pearson's correlation coefficient (R) at 0.70. None of the predictors were highly correlated.

Multiple (backward) logistic regression analyses were performed (removal 0.05) to determine which baseline factors were associated with 1) consultation with a specialist and 2) surgery. The analyses with the outcome consultation were performed with all candidate predictors. The analyses with the outcome surgery were performed with a restricted number (N = 17) of candidate predictors to ensure sufficient power. The restricted candidate predictors included were: age, gender, body mass index, attitude and beliefs about low back pain (BBQ), duration of back symptoms, severity of back pain, presence of radiating pain in the legs below the knee, severity of leg pain, neurological symptoms of the legs, presence of continuous back pain independent of posture or activity, disability (RDQ), history of back surgery, disc herniation, nerve root compression, spinal stenosis, spondylolisthesis, and serious pathology. If potential predictors were selected in at least 3 of 5 imputed databases in the multivariate analysis, they were included in the final model.

To evaluate the discriminative ability of the models, a receiver operating characteristic curve was generated for the predicted probabilities and the area under the curve (AUC) was calculated [15]. Analyses were performed using SPSS version 20.0 (SPSS Inc., USA).

Results

A total of 683 referred patients participated in the study (Figure 1). During follow-up, 547 (80%) patients returned the 3-month follow-up questionnaire and 474 (69%) patients returned the 12-month follow-up questionnaire. Information on BMI at baseline was missing in 8 patients (1%); information on severity of leg pain (NRS) at baseline was missing in

Table 1 Baseline characteristics of the included 683 patients

	Study population *n* = 683	**Consultation specialist* *n* = 301**	**Surgery** *n* = 124**
Patient characteristics			
Age in years, mean (SD)	49.9 (12.5)	50.6 (12.9)	49.5 (14.0)
Male	365 (53)	157 (52)	71 (57)
BMI, mean (SD)	25.9 (3.8)	26.0 (3.9)	26.3 (3.8)
Education level low	244 (36)	105 (35)	45 (36)
Employed (paid job)	479 (70)	204 (68)	93 (75)
Attitude and beliefs about back pain (BBQ), mean (SD)	26.3 (6.1)	26.4 (6.1)	26.5 (6.1)
Back pain characteristics			
History of back pain	549 (80)	239 (79)	96 (77)
Acute back pain (<3 months)	228 (33)	108 (36)	41 (33)
Severity of back pain (NRS), mean (SD)	6.6 (2.0)	6.8 (2.0)	7.0 (1.8)
Pain radiating in the leg below the knee	450 (66)	209 (69)	101 (82)
Severity of leg pain (NRS), mean (SD)	5.6 (2.8)	6.1 (2.7)	6.6 (2.4)
Neurological symptoms legs	525 (77)	245 (81)	112 (90)
Morning stiffness of the back	353 (52)	159 (53)	62 (50)
Continuous back pain	347 (51)	164 (55)	68 (55)
Disability (RDQ), mean (SD)	13.5 (5.2)	14.5 (4.9)	15.2 (4.6)
History of back surgery	112 (16)	66 (22)	32 (26)
Recent back surgery (< 1 year)	22 (3)	10 (3)	6 (5)
Recent consultation specialist (< 3 months)	107 (16)	61 (20)	23 (19)
MRI findings			
Bulging	308 (45)	136 (45)	54 (44)
Disc herniation	492 (72)	212 (70)	101 (82)
Nerve root compression	472 (69)	216 (72)	109 (88)
Spinal stenosis	87 (13)	45 (15)	28 (23)
Spondylolisthesis	56 (8)	30 (10)	15 (12)
Serious pathology (fracture, malignancy or discitis)	22 (3)	11 (4)	5 (4)

Data are presented as numbers (percentages) unless otherwise indicated
* Consultation with a specialist (orthopaedic, neurologist, neurosurgeon, or rheumatologist) during the first 3 months of follow-up
** Surgery during 12 months follow-up
SD: standard deviation; BBQ: Back Beliefs Questionnaire (range 9-45), a higher score indicates better attitude/beliefs regarding back pain; NRS: numeric rating scale (range 0-10, 0 means no pain); RDQ: Roland Disability Questionnaire (range 0-24), a higher score indicates worse health

2 patients (0.4%); and information on consultation with specialists prior to inclusion was missing in 1 patient (0.1%).
The baseline characteristics of the patients are presented in Table 1. The mean age of the patients was 49.9 (SD 12.5; range 19-80 years). In total, 53% of the patients were male, 36% had a low education level and 70% had a paid job. At baseline, the mean back pain severity was 6.6 (SD 2.0) and 33% of the patients reported acute back pain. Of all patients, 66% reported radiating pain in the leg below the knee; 77% reported neurological symptoms of the leg(s). The mean disability score (RDQ) was 13.5 (SD 5.2). The MRI reports described the presence of disc herniation in 72% of the patients; 69% of the MRI reports mentioned signs of nerve root compression. Spinal stenosis was reported in 13% of the patients and spondylolisthesis was reported in 8% of the patients. Serious pathologies (fractures, malignancies and discitis) were reported in 3% of the patients.

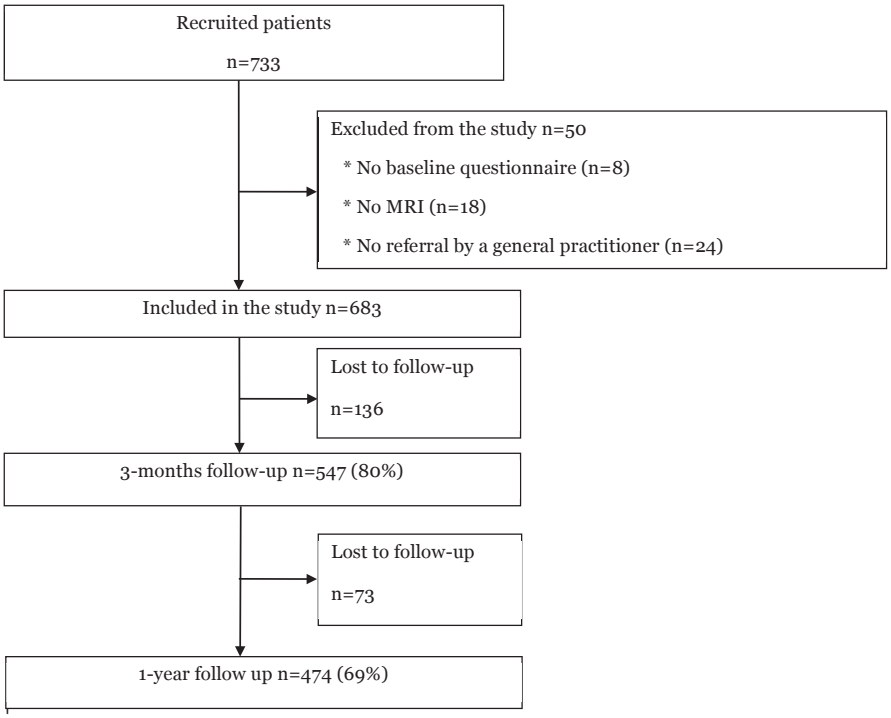

Figure 1. Flow chart of the study population

Healthcare use

More than half of the patients (301 of 547, 55%) reported consultation with a specialist (orthopedic surgeon, neurologist, neurosurgeon, or rheumatologist) during the first 3 months of follow-up (Table 2). Consultations with a psychologist or multidisciplinary pain team were rarely reported by the patients. Consultation with the general practitioner decreased after 3 months of follow-up to 65% of the patients (354 of 547), and after 12 months of follow-up to 21% (100 of 474 patients). At baseline, 79% of the patients reported consultation with a physiotherapist during the 3 months prior to the baseline measurement. After 12 months follow-up this had decreased to 47% (221 of 474).

Regarding medication use, 559 (82%) of all patients took pain medication for their back pain at baseline. More than one analgesic was consumed by 337 (49%) patients. The most often used analgesics were NSAIDs (63%).

Table 2 Healthcare use* at baseline, and at 3 and 12-month follow-up.

	Baseline (*n* = 683)		**3-month follow-up (*n* = 547)**		**12-month follow-up (*n* = 474)**	
Consultation						
General practitioner	-		354	(65)	100	(21)
Physiotherapist	536	(79)	417	(76)	221	(47)
Specialist**	107	(16)	301	(55)	97	(21)
Company doctor	88	(13)	143	(26)	58	(12)
Psychologist	14	(2)	23	(4)	21	(4)
Multidisciplinary pain team	7	(1)	13	(2)	12	(3)
Medication use						
Medication use	559	(82)	373	(68)	185	(39)
Paracetamol	267	(39)	192	(35)	89	(19)
NSAIDs	429	(63)	235	(43)	117	(25)
Tramadol	126	(18)	77	(14)	27	(6)
Benzodiazepines	82	(12)	34	(6)	15	(3)
Strong opioids	65	(10)	51	(9)	9	(2)
Antidepressants/ Anticonvulsants	23	(3)	22	(4)	11	(2)
Other	4	(1)	0	(0)	2	(0.4)
Unknown	19	(3)	13	(2)	3	(1)

Data are presented as numbers (percentages)

* During the last 3 months

** Includes orthopaedic surgeon, neurologist, neurosurgeon, and rheumatologist

Table 3 Results of multivariate logistic regression analysis regarding potential baseline predictors for consultation with a specialist* (n = 547)

Consultation specialist	Pooled OR (95% CI)		P-value	AUC
				0.70
Severity of leg pain (NRS)	1.1	(1.0 - 1.2)	< 0.01	
Disability (RDQ)	1.1	(1.0 - 1.1)	< 0.01	
History of back surgery (yes)	2.3	(1.3 - 3.8)	< 0.01	
Consultation specialist before baseline** (yes)	2.1	(1.2 - 3.6)	< 0.01	
Spinal stenosis on MRI (yes)	2.3	(1.3 - 4.2)	< 0.01	

* Consultation with a specialist (orthopaedic, neurologist, neurosurgeon, or rheumatologist) during the first 3 months of follow-up
** Consultation with a specialist in the 3 months prior to the baseline measurement
OR: odds ratio; CI: confidence interval; AUC: area under the curve; NRS: numeric rating scale (range 0-10, 0 means no pain); RDQ: Roland Disability Questionnaire (range 0-24), a higher score indicates worse health

Strong opioids were used in 65 (10%) patients. Usage of neuropathic pain medication (antidepressants or anticonvulsants) were rarely reported. At 12 months follow-up, the number of patients taking pain medication decreased to 185 (39%) patients.

Predictors of health care use

Table 3 shows the results of the multivariate logistic regression analysis regarding the potential predictors on the outcome consultation with a specialist in the first 3 months of follow-up. In the final model the variables associated with consultation with a specialist were: severity of leg pain (NRS) [odds ratio (OR) 1.1; 95% confidence interval (CI) 1.0-1.2], disability (RDQ) (OR 1.1; CI 1.0-1.1), history of back surgery (OR 2.3; CI 1.3-3.8), consultation with a specialist before baseline (OR 2.1; CI 1.2-3.6) and spinal stenosis seen on MRI (OR 2.3; CI 1.3-4.2). The AUC for the final model was 0.70.

During the 12 months follow-up, 124 patients (18%) underwent surgery for their low back pain. Table 4 shows the results of the multivariate logistic regression analysis regarding the potential predictors on the outcome surgery. In the final model the variables associated with surgery during 12 months follow-up were: age (OR 0.98; CI 0.96-1.0), pain radiating in the leg below the knee (OR 1.9; CI 1.1-3.3), baseline disability score on RDQ (OR 1.1; CI 1.0-1.1), history of back surgery (OR 2.1; CI 1.2-3.7), nerve root compression (OR 2.8; CI 1.5-5.2) and spinal stenosis seen on MRI (OR 3.2; CI 1.7-6.0). The AUC for the final model was 0.75.

Table 4 Results of multivariate logistic regression analysis regarding potential baseline predictors for surgery during 12-month follow-up (n = 477)

Surgery	Pooled OR (95% CI)		P-value	AUC
				0.75
Age	0.98	(0.96 - 1.0)	< 0.05	
Pain radiating in the leg below the knee (yes)	1.9	(1.1 - 3.3)	< 0.01	
Disability (RDQ)	1.1	(1.0 - 1.1)	< 0.01	
History of back surgery (yes)	2.1	(1.2 - 3.7)	< 0.01	
Nerve root compression on MRI (yes)	2.8	(1.5 - 5.2)	< 0.01	
Spinal stenosis on MRI (yes)	3.2	(1.7 - 6.0)	< 0.01	

OR: odds ratio; CI: confidence interval; AUC: area under the curve; RDQ: Roland Disability Questionnaire (range 0-24), a higher score indicates worse health

Discussion

This study presents the health care use of the 683 included patients who were referred for MRI of the lumbar spine by their general practitioner and identified predictors for consultation with a specialist and surgery during follow-up. Patients were frequently seen by a specialist during the first 3 months of follow-up (301 to 547; 55%), and 124 (18%) patients underwent spine surgery during the 12 month follow-up. Five baseline characteristics were associated with consultation, including four characteristics from history taking. Younger patients, presence of pain radiating in the leg below the knee, higher baseline disability score, recurrent back surgery, presence of nerve root compression and/or spinal stenosis on MRI were predictive for subsequent spine surgery during the 12 months follow-up.

Reported back pain severity at baseline was higher than reported in previous low back pain cohort studies [16-19]. Severity of disability resemble reported data in two studies that included low back pain patients referred for MRI or radiography by their primary physician [17, 20]. The MRI reports showed disc herniation in 72%, and nerve root compression in 69% of the patients. The prevalence of disc herniation was higher than reported in the study of You et al. including patients referred for MRI in primary care [7]. On the contrary, the prevalence of spinal stenosis was substantially lower (13% vs. 48%). The presence of spinal stenosis could be underestimated as a consequence of using the MRI reports instead of standardized scoring of the MR images. The most common serious pathol-

ogy observed was vertebral fracture (3%). This is consistent with a recent study about the prevalence of serious spinal pathology in primary care [21].

Predictors of health care use

In this study, patients with higher leg pain scores, higher score for disability, a history of back surgery, a recent consultation with a specialist or spinal stenosis seen on MRI were more likely to consult a specialist during the first 3 months of follow-up (AUC 0.70). Spinal stenosis was the only MRI finding that predicted referral to a specialist. These findings only partially correspond with the conclusions of the Canadian study of You et al [7], which reported that patients with disc herniation or spinal stenosis on MRI were predictive of subsequent consultation with a spine surgeon, with likelihood ratios of 2.01 and 1.52 respectively. This discrepancy can possibly be explained by the difference between the Dutch and Canadian health care system.

In the Netherlands, general practitioners have a gatekeeping role in the health care system which implies that patients cannot consult a hospital specialist without a referral from their general practitioner. To maintain this gatekeeping role, the Dutch general practitioners tend to adhere strictly to their clinical guidelines formulated by the Dutch College of General Practitioners [3]. These guidelines recommend that subsequent consultation with a spine surgeon should not be performed in all patients with disc herniation or nerve root compression seen on MRI, but only in patients who have chronic, severe radiculopathy (i.e. those who may be surgical candidates). It could be that because of this, disc herniation or nerve root compression seen on MRI was not a significant predictor.

The AUC for the final model was moderate (0.70). A possible reason for this could be that there may be other determinants of general practitioners' decisions to refer, but not included in this study, like the patients request for a referral even if this is not indicated according to the guidelines [22]. This is supported by a recent survey of Canadian primary care physicians, which noted that the most common reasons primary care physicians refer to a spine surgeon were not only compression of neurological structures reported on imaging or persistent symptoms but also patients request for a consultation [23]. Future studies need to include this possible determinant of general practitioners' decisions to refer.

The decision to perform surgery is made after the referral by the general practitioner, mostly by the neurosurgeon together with the patient. In our study, 124 patients (18%) received spinal surgery. Patients with low age, pain radiating in the leg below the knee, high disability, a history of back surgery, nerve root compression or spinal stenosis were more likely to undergo subsequent spine surgery (AUC 0.75). These findings largely correspond with the conclusions of the study of Cheng et al. [8]. In this retrospective study of 1586 symptomatic patients from a single academic spine surgery practice, MRI findings of disc herniation, spinal stenosis, and spondylolisthesis were independent predictors of surgical candidacy. The findings also partially correspond with the conclusions of the study of Peul et al. [9], which reported high leg pain and more disability at baseline as prognostic factors for subsequent surgery in sciatica patients.

Our study had several strengths. It included a relatively high number of patients, it had low drop-out rates despite the use of online questionnaires, and it had almost no missing data (ranging from 0.1% to 1%). However, there are also some limitations of our study that need to be considered when interpreting the results. One limitation is that the presence of several MRI findings could be underestimated, such as spinal stenosis, as a consequence of using the MRI reports instead of standardized scoring of the MR images. However, by using the MRI reports we reflect daily general practice as close as possible.

Conclusions

At 3 months follow-up, 55% of patients with low back pain referred for MRI in general practice reported consultation with a specialist. Five baseline characteristics were associated with consultation; including four characteristics from history taking. During 12 months of follow-up 18% of the patients underwent surgery. Six clinic baseline characteristics were associated with surgery during 12 months follow-up; including nerve root compression and spinal stenosis seen on MRI.

References

1. Chou, R., R.A. Deyo, and J.G. Jarvik, Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am*, 2012. 50(4): p. 569-85.
2. Williams, C.M., et al., Low back pain and best practice care: A survey of general practice physicians. *Arch Intern Med*, 2010. 170(3): p. 271-7.
3. Chavannes AW, M.J., Koes BW, Lubbers WJ, Ostelo R, Spinnewijn WEM, Kolnaar BGM, NHG-Standaard Aspecifieke lagerugpijn. *Huisarts Wet*, 2005. 48(3): p. 113-123.
4. Chou, R., et al., Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet*, 2009. 373(9662): p. 463-72.
5. Chou, R., et al., 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*, 2007. 147(7): p. 478-91.
6. van Tulder, M., et al., Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J*, 2006. 15 Suppl 2: p. S169-91.
7. You, J.J., et al., Patterns of care after magnetic resonance imaging of the spine in primary care. *Spine (Phila Pa 1976)*, 2013. 38(1): p. 51-9.
8. Cheng, F., J. You, and Y.R. Rampersaud, Relationship between spinal magnetic resonance imaging findings and candidacy for spinal surgery. *Can Fam Physician*, 2010. 56(9): p. e323-30.
9. Peul, W.C., et al., Improving prediction of "inevitable" surgery during non-surgical treatment of sciatica. *Pain*, 2008. 138(3): p. 571-6.
10. Valls, I., et al., Factors predicting radical treatment after in-hospital conservative management of disk-related sciatica. *Joint Bone Spine*, 2001. 68(1): p. 50-8.
11. Verwoerd, A.J., et al., Systematic review of prognostic factors predicting outcome in non-surgically treated patients with sciatica. *Eur J Pain*, 2013.
12. Vroomen, P.C., M.C. de Krom, and J.A. Knottnerus, When does the patient with a disc herniation undergo lumbosacral discectomy? *J Neurol Neurosurg Psychiatry*, 2000. 68(1): p. 75-9.
13. Symonds, T.L., et al., Do attitudes and beliefs influence work loss due to low back trouble? *Occup Med (Lond)*, 1996. 46(1): p. 25-32.
14. Moons, K.G., et al., Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol*, 2006. 59(10): p. 1092-101.
15. Harrell, F.E., Jr., Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. 2001, New York: Springer.

16. Campbell, P., et al., Prognostic Indicators of Low Back Pain in Primary Care: Five-Year Prospective Study. *J Pain*, 2013.
17. Graves, J.M., et al., Early imaging for acute low back pain: one-year health and disability outcomes among Washington State workers. *Spine (Phila Pa 1976)*, 2012. 37(18): p. 1617-27.
18. Hill, J.C., et al., Clinical outcomes among low back pain consulters with referred leg pain in primary care. *Spine (Phila Pa 1976)*, 2011. 36(25): p. 2168-75.
19. Mehling, W.E., et al., The prognosis of acute low back pain in primary care in the United States: a 2-year prospective cohort study. *Spine (Phila Pa 1976)*, 2012. 37(8): p. 678-84.
20. Jarvik, J.G., et al., Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *JAMA*, 2003. 289(21): p. 2810-8.
21. Henschke, N., et al., Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. *Arthritis Rheum*, 2009. 60(10): p. 3072-80.
22. Bederman, S.S., et al., Referral practices for spinal surgery are poorly predicted by clinical guidelines and opinions of primary care physicians. *Med Care*, 2010. 48(9): p. 852-8.
23. Rampersaud YR, A.J., Harvey B, et al. , Understanding primary care physicians' challenges, barriers, and priorities in caring for patients with low back pain. *Can J Surg*, 2011. 54: p. S4.

5

Diagnosis of lumbar spinal stenosis: an updated systematic review of the accuracy of diagnostic tests

Evelien I.T. de Schepper,
Gijsbert M. Overdevest, Pradeep Suri,
Wilco C. Peul, Edwin H.G. Oei,
Bart W. Koes, Sita M.A. Bierma-Zeinstra,
Pim A.J. Luijsterburg

Abstract

Study Design

Systematic review of diagnostic studies.

Summary of Background Data

A wide range of clinical, radiologic and electrodiagnostic tests are used to diagnose lumbar spinal stenosis. An accurate diagnosis is vital, because lumbar spinal stenosis may require specific medical advice and treatment. Therefore, it is important to know the accuracy of these diagnostic tests currently available.

Objective

To update our previous systematic review on the diagnostic accuracy of tests used to diagnose lumbar spinal stenosis.

Methods

A comprehensive literature search was conducted for original diagnostic studies on lumbar spinal stenosis, in which one or more diagnostic tests were evaluated with a reference standard, and diagnostic accuracy was reported or could be calculated. Our previous systematic review included studies up to March 2004; this review is current up to March 2011. Included studies were assessed for their methodological quality using the Quadas tool. Study characteristics and reported diagnostic accuracy were extracted.

Results

Twenty-two additional articles over the 24 included in the previous review met the inclusion criteria. Combined, this resulted in twenty articles concerning imaging tests, 11 articles evaluating electrodiagnostic tests, and 15 articles evaluating clinical tests. Estimates of the diagnostic accuracy of the tests differed considerably.

Conclusions

There is a need for a consensus on criteria to define and classify lumbar spinal stenosis. At present the most promising imaging test for lumbar spinal stenosis is MRI, avoiding myelography because of its invasiveness and lack of superior accuracy. Electrodiagnostic studies showed no superior accuracy for conventional electrodiagnostic testing compared to MRI. These tests should be considered in the context of those presenting symptoms with the highest diagnostic value, including radiating leg pain that is exacerbated while standing up, the absence of pain when seated, the improvement of symptoms when bending forward, and a wide-based gait.

Introduction

Lumbar spinal stenosis (LSS) is commonly used to describe patients with symptoms related to an anatomic reduction of the lumbar spinal canal size [1]. The challenge to the anatomically based definition is that while necessary for the diagnosis of LSS, it is not sufficient to determine the severity of symptoms that leads a patient to seek treatment [1]. The extent of narrowing of the spinal canal correlates poorly with symptom severity and radiologically significant lumbar stenosis can be found in asymptomatic individuals [1-4]. Furthermore, lower extremity pain, numbness, or weakness is frequently seen in the setting of low back pain, and other causes abound. As a consequence, correlating symptoms and physical examination findings with imaging results is necessary to establish a definitive diagnosis [1]. Unfortunately, there is no generally accepted gold standard for the diagnosis of LSS [5, 6]. A wide range of clinical, electrodiagnostic, and radiologic tests are currently used to diagnose LSS. It is important to know the diagnostic value of these tests because false-positive test results may lead to unnecessary surgery and/or expensive or invasive additional diagnostic interventions.
Prior studies as recent as 2006 have concluded that no firm conclusions could be drawn regarding diagnostic accuracy of different tests due to poor study quality [7, 8]. New diagnostic studies have since been published, with more recently developed diagnostics tests, and possibly with increasing study design quality.

In this article, we performed an update of our previous systematic review [8] and systematically reviewed the diagnostic accuracy of tests for the assessment of LSS.

Materials and methods

Data Sources and Searches
All 24 articles considered in the previous review [8] were directly included in the present one. The previous review was updated up to March 2004. An additional literature search using the same search strategy and restricted to March 2004 up to March 2011 was performed in Medline (Pubmed) and Embase.

Study Selection

The following selection criteria were used:

1) The study investigated the diagnostic accuracy of imaging, clinical examination, and other tests in detecting lumbar spinal stenosis, in an adult study population, 2) one or more different diagnostic tests as well as a reference test were included within the design, 3) diagnostic accuracy was reported or could be calculated, 4) if the results concerned a subgroup of patients with lumbar spinal stenosis, these were analyzed separately in the same article, 5) the article was written in English, German, French, or Dutch.

For this update, two reviewers read all titles/abstracts, independently of each other. Articles that could not be excluded based on title and/or abstract were retrieved in full text and were read and checked for inclusion by two reviewers independently. If there was no agreement, a third reviewer made the final decision.

Additionally, reference lists of all included articles were reviewed to search for additional relevant articles.

Data Extraction and Quality Assessment

Two reviewers independently extracted the data. Data describing study design, characteristics of the study population, test characteristics and diagnostic parameters were extracted. In order to gain insight in the diagnostic accuracy, we focused on the sensitivity and the specificity of the test at issue.

Four independent reviewers assessed the risk of bias of each included study using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool [9]. The QUADAS tool consists of 11 items that refer to internal validity. We added 4 items related to the criteria to diagnose LSS, the interobserver variation and the index test (Appendix Figure 1). All 22 studies of the original review were additionally scored for the additional 4 items by one author. A radiologist was consulted for the assessment of the used technology of the index test (item 12). Disagreements were resolved by consensus and in case of persisting disagreement a third review author was consulted. We did not use a summary score since the interpretation of summary score is problematic and potentially misleading [9, 10].

Data Synthesis and Analysis

All reported calculations and results in the studies were checked. When the diagnostic outcomes were not reported, we calculated them if sufficient

data were presented. The confidence intervals (CI) of the sensitivity and specificity were also calculated. When the sensitivity and specificity were not reported and could not be calculated, we extracted other values if possible such as positive predictive value (PPV) and negative predictive value (NPV). Because of the heterogeneity of the tests, study population, and reference standards, statistical pooling was not possible. Therefore, the results are summarized in a qualitative manner.

Reference standard

A diagnosis of the clinical syndrome of LSS requires both the presence of characteristic symptoms and signs and radiographic or anatomic confirmation of narrowing of the lumbar spinal canal [11]. In this update, we clearly made the distinction between studies using a clinical reference standard and studies using an anatomic reference standard. A clinical reference standard was defined when it included expert opinion based on clinical findings and imaging and/or surgery, and the spectrum of patients was representative of the patients who would receive the test in clinical practice. Studies using an anatomic reference standard use only imaging and/or surgery findings to diagnose LSS.

Results

Search and selection

In this update, our search strategy in Medline (Pubmed) resulted in 714 references and Embase yielded an additional 85 references. Reviewing of the reference lists resulted in 19 additional articles. In total, 63 articles were retrieved in full text. Twenty-two of these articles met the inclusion criteria. Including the 24 articles from our previous review, a total of 46 articles were included for this systematic review [12-58] (Figure 1). Main reasons for exclusion were: lack of reference standard, diagnostic accuracy was not reported or could not be calculated, study design was a case report or case series; inclusion of cervical or thoracic stenosis cases; and/or no separate outcomes for cases of LSS were reported.

Type of Studies

Twenty articles evaluated the diagnostic accuracy of imaging tests (i.e., CT, MRI, myelography and ultrasound), 11 articles (describing 7 study populations) evaluated electrodiagnostic tests (i.e., electromyography, dermatomal somatosensory-evoked potentials and caudal motor conduction time),

and 15 articles (describing 12 study populations) evaluated clinical tests (e.g., standardized history, physical examination, pain drawings and gait analyses).
The characteristics of the included studies on imaging tests (N = 20) are shown in Appendix Table 1, on electrodiagnostic tests (N = 7) in Appendix Table 2, and on clinical tests (N = 12) in Appendix Table 3.

Quality Assessment
The risk of bias assessment of the individual studies is presented in Appendix Figure 2. The initial agreement between the reviewers was 75% for imaging studies, 77% for electrodiagnostic studies and 87% for clinical studies. The initial disagreements were resolved by consensus.

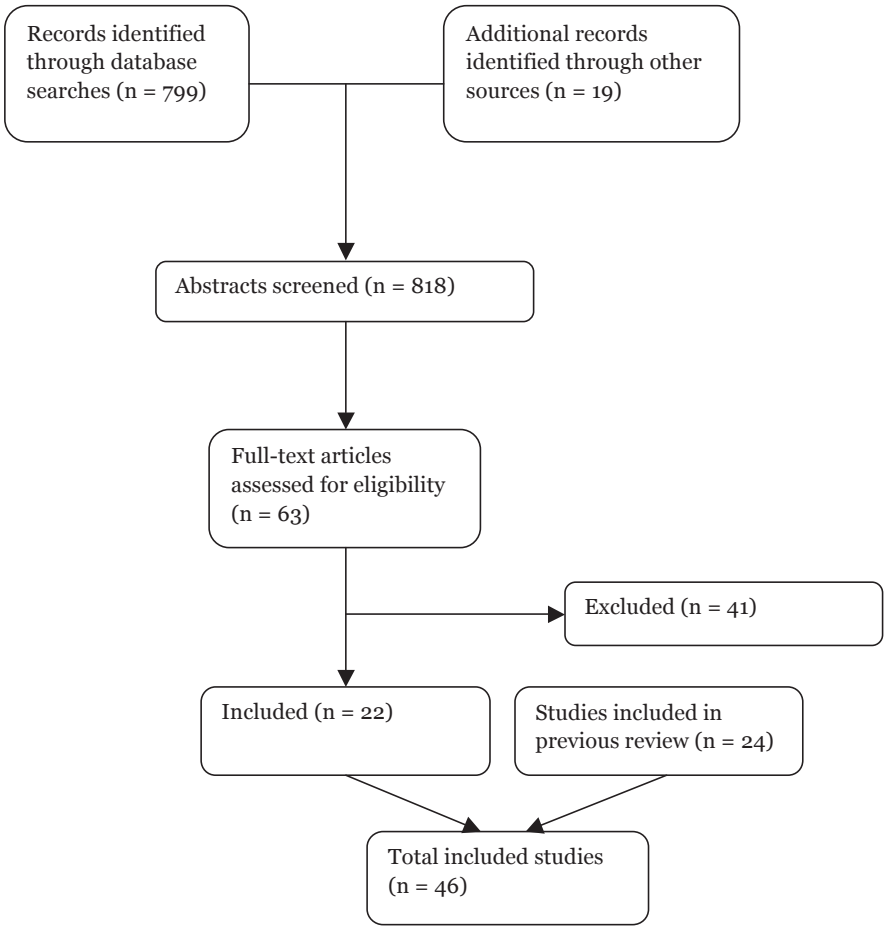

Figure 1. Flow chart

Almost none of the studies reported the time period between the index test and reference test (item 3), or the interobserver variation of the interpretation of test results (item 15). The majority of the imaging studies results may be influenced by knowledge of the results of the reference standard (test review bias). On the other hand, incorporation bias was avoided in the majority of the recent studies (item 6). Overall, studies with a more recent date of publication tended to have less bias.

Imaging tests

Table 1 presents data on the diagnostic accuracy of the imaging tests. All imaging studies used an anatomic reference standard. Summaries of sensitivity and specificity for each imaging test (MRI, CT, etc…) are provided as online content (see Document, Supplemental Digital Content 1).

Overall The studies showed no superior accuracy for myelography compared to CT, MRI, or 3D-MRM, with MRI and 3D-MRM showing the highest sensitivity. Three-sequence MRI appeared to be more sensitive than single-sequence MRI. The accuracy of ultrasound appeared to be almost equal to that of CT or myelography. The results of a single study about the nerve root sedimentation sign suggested a sensitivity of 94% and a specificity of 100%.

Electrodiagnostic tests

Table 2 presents data on the diagnostic accuracy of the electrodiagnostic tests. This review is updated with 5 additional electrodiagnostic studies [21, 28, 29, 32-35, 52, 55]. The electrodiagnostic tests consisted of conventional electromyography (EMG) and nerve conduction studies (NCS), dermatomal somatosensory-evoked potentials (DSEP), and the assessment of caudal motor conduction time (CMCT) with magnetic stimulation. Four separate articles [21, 32-34] described one study population, but varied in study design and used different reference standards. The reference standards were: expert opinion based on a combination of clinical, radiologic, and other diagnostic tests; MRI or CT; surgery and myelography. Summaries of sensitivity and specificity for each electrodiagnostic test (EMG, NCS, etc…) are provided as online content (see Document, Supplemental Digital Content 2).

Overall The diagnostic accuracy of electrodiagnostic testing was only modest for most electrodiagnostic tests studied. Paraspinal mapping had a high specificity in two studies and may increase the likelihood of LSS when using a reference standard of expert opinion based on clinical and

Table 1 Diagnostic Accuracy of Imaging Tests for Lumbar Spinal Stenosis

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
		MRI			
Aota 2007 [13]	117	0.96 (0.88-1.00)	0.68 (0.65-0.71)	3.0	0.06
Barz 2010 [15]	200				
Nerve root sedimentation sign		0.94 (0.89-0.99)	1.00	~	0.06
Bischoff 1993 [17]	28	0.87 (0.73-0.96)	0.75 (0.35-0.97)	3.5	0.17
Chang 2010 [20]	13				
Root-lift-up sign		0.92 (0.82-0.99)	0.60 (0.48-0.72)	2.3	0.15
Sagittal sign		0.77 (0.65-0.90)	0.43 (0.31-0.56)	1.4	0.52
Jia 1991 [39]	27	0.85 (0.66-0.96)	-	-	-
Modic 1986 [48]	48	0.77 (0.58-0.90)	0.88 (0.71-0.96)	6.4	0.26
Rankine 1997 [50]	79	0.60 (0.32-0.84)	0.95 (0.90-1.00)	12	0.42
Yan 2010 [56]	22	0.88 (0.73-1.00)	0.60 (0.17-1.00)	2.2	0.20
		3D-MR Myelography			
Aota 2007 [13]	117				
Abnormal course of nerve roots		0.96 (0.88-1.00)	0.84 (0.82-0.86)	6.0	0.05
Spinal nerve swelling		0.60 (0.41-0.79)	0.99 (0.98-1.00)	55	0.40
Eberhardt 1994 [25]	65	1.00 (0.96-1.00)	-	-	-
Freund 1997 [30]	25	1.00 (0.86-1.00)	-	-	-
		CT			
Arrault 1987 [14]	60			-	-
Central stenosis		0.72 (0.55-0.86)	-	-	-
Lateral stenosis		0.62 (0.48-0.75)	-	-	-
Isolated bony lateral stenosis		0.57 (0.34-0.78)	-	-	-
Bell 1984 [16]	46				
Strong diagnostic criteria		0.50	-	-	-

Firm diagnostic criteria		0.74	-	-	-
Bolender 1985 [18]	24				
AP diameter		0.21 (0.07-0.42)	-	-	-
Cross sectional area		0.92 (0.73-1.00)	-	-	-
Donmez 1990 [23]	50				
AP diameter, central stenosis		0.89 (0.65-0.99)	0.81 (0.65-0.91)	4.5	0.14
Interpedincular distance		0.72 (0.47-0.90)	0.98 (0.87-1.00)	36	0.29
Lateral recess stenosis		0.79 (0.60-0.92)	0.95 (0.82-0.99)	16	0.22
Feldmeyer 1982 [27]	7	1.00 (0.59-1.00)	-	-	-
Modic 1986 [48]	48	0.79 (0.58-0.93)	0.90 (0.73-0.98)	7.9	0.23
Yan 2010 [56]	22	0.76 (0.56-0.97)	0.60 (0.17-1.00)	1.9	0.39
		Myelography			
Arrault 1987 [14]	60			-	-
Central stenosis		0.56 (0.38-0.72)	-	-	-
Lateral stenosis		0.62 (0.48-0.75)	-	-	-
Isolated bony lateral stenosis		0.62 (0.38-0.82)	-	-	-
Bell 1984 [16]	46			-	-
Strong diagnostic criteria		0.67	-	-	-
Firm diagnostic criteria		0.87	-	-	-
Bischoff 1993 [17]	28	0.82 (0.66-0.92)	0.88 (0.47-1.00)	6.8	0.20
Bolender 1985 [18]	24	0.92 (0.73-1.00)	-	-	-
Eberhardt 1994 [25]	65	0.71 (0.60-0.80)	-	-	-
Feldmeyer 1982 [27]	7	0.71 (0.29-0.96)	-	-	-
Freund 1997 [30]	25	1.00 (0.86-1.00)	-	-	-
Herkowitz 1982 [36]	18	0.94 (0.73-1.00)	1.00	-	-
Jia 1991 [39]	27	0.90	-	-	-
Modic 1986 [48]	48	0.54 (0.33-0.74)	0.91 (0.75-0.98)	6	0.5

CT-myelography					
Bischoff 1993 [17]	28	0.87 (0.73-0.96)	0.75 (0.35-0.97)	3.5	0.17
Epidurography					
Yan 2010 [56]	22	0.94 (0.83-1.00)	0.80 (0.45-1.00)	4.7	0.07
Epidural venography					
Herkowitz 1982 [36]	18	0.77 (0.52-0.94)	1.00	~	0.23
Radiography					
Ilkko 1989 [37]	116				
Short pedicles		0.55 (0.36-0.74)	0.85 (0.76-0.92)	3.7	0.53
High narrow intervertebral foramina		0.52 (0.33-0.71)	0.84 (0.74-0.91)	3.2	0.58
Thick lamina's		0.21 (0.80-0.40)	0.94 (0.87-0.99)	3.6	0.84
Sagittal intervertebral joints		0.08 (0.01-0.26)	0.94 (0.86-0.98)	1.3	0.98
Small interlaminar window		0.38 (0.21-0.58)	0.84 (0.74-0.91)	2.4	0.74
Deep posterior concavity of vertebral bodies		0.24 (0.10-0.44)	0.87 (0.79-0.94	1.8	0.86
Three or more criteria fulfilled		0.66 (0.46-0.82)	0.93 (0.86-0.97)	9.4	0.37
Ultrasound					
Engel 1985 [26]	19	0.95 (0.74-1.00)	1.00	~	0.05
Tervonen 1989 [54]	76	0.90 (0.55-1.00)	0.96 (0.89-1.00)	23	0.10
Ultrasound calcaneus					
Mariconda 2004 [47]	117				
SOS (cut-off 1535.5)		0.52 (0.39-0.64)	0.70 (0.57-0.82)	1.7	0.69
Over 60 years, SOS (cut-off 1532)		0.51	0.80	2.6	0.61
Over 60 years, BUA (cut-off 125.5)		0.51	0.84	3.2	0.58
Over 60 years, stiffness (cut-off 89.5)		0.59	0.72	2.1	0.57
Men SOS (cut-off 1535.5)		0.79 (0.65-0.94)	0.65 (0.46-0.85)	2.3	0.32
Men BUA (cut-off 125.5)		0.83 (0.69-0.97)	0.57 (0.36-0.77)	1.9	0.31
Men stiffness (cut-off 95)		0.79 (0.65-0.94)	0.61 (0.41-0.81)	2.0	0.34
Men over 60 years, SOS (cut-off 1532)		0.89	0.75	3.6	0.15
Men over 60 years, BUA (cut-off 125.5)		0.89	0.67	2.7	0.16
Men over 60 years, stiffness (cut-off 96.5)		0.83	0.75	3.3	0.23

radiologic data [33, 55]. This method therefore may have some utility in confirming the clinical significance of radiological LSS among subjects with atypical symptoms. The diagnostic accuracy of DSEP and Magnetic stimulation MCT remains unclear.

Clinical tests

Table 3 presents data on the diagnostic accuracy of the clinical tests. This review is updated with 7 additional clinical studies [12, 22, 40, 42, 43, 49, 57, 58]. One study had many aspects that were unclear or internally contradictory, therefore no data are reported [58].

The clinical tests consisted of questionnaires, standardized history and physical examination, gait-analyses, treadmill tests, and pain drawings analyzed in three different ways but in the same study population [44-46]. The reference standards were: expert opinion based on a combination of clinical, radiologic, and other diagnostic tests; MRI or CT; fluoroscopically guided injections and myelography. One population was studied in 2 separate reports: the first involved history and physical examination findings, and the second, questionnaire items [42, 57].

Summaries of sensitivity and specificity for each clinical test are provided as online content (see Document, Supplemental Digital Content 3).

Overall The symptoms of radiating leg pain, thigh pain and pain that is exacerbated while standing up showed the highest sensitivity for LSS. Bilateral buttock or leg pain, the absence of pain when seated, the improvement of symptoms when bending forward, and a wide-based gait were generally the most useful clinical findings for ruling in the diagnosis of LSS, as reflected by large magnitude likelihood ratios (> 5.0 or < 0.20), while having at least fair moderate sensitivity [12, 41, 42]. In contrast, the clinical findings of symptoms related to cauda equina syndrome and urinary disturbances were highly specific, but insensitive [42]. In general, individual physical examination tests were not as useful as symptoms. Simple clinical diagnostic support tools may help to synthesize the independent diagnostic value of combinations of history and physical examination measures.

Table 2 Diagnostic Accuracy of Electrodiagnostic Tests for Lumbar Spinal Stenosis

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
Electromyography and Nerve Conduction					
Chiodo 2007 [21]	32				
MRI		-	0.44 (0.27-0.61)	-	-
Electrodiagnosis		-	0.59 (0.42-0.76)	-	-
Fisher 2007, 2008 [28,29]	21				
Needle EMG		0.60 (0.30-0.90)	0.82 (0.59-1.00)	3.3	0.49
Needle EMG and NCS		0.90 (0.71-1.00)	0.45 (0.16-0.75)	1.7	0.22
Any abnormality (NC-stat)		0.90 (0.71-1.00)	0.27 (0.01-0.54)	1.2	0.37
Any mean F-wave abnormality (NC-stat)		0.80 (0.55-1.00)	0.55 (0.25-0.84)	1.8	0.37
Any CMAP abnormality (NC-stat)		0.60 (0.30-0.90)	0.82 (0.59-1.00)	3.3	0.49
Any mean F-wave or CMAP abnormality (NC-stat)		0.80 (0.55-1.00)	0.55 (0.25-0.84)	1.8	0.37
Any peroneal abnormality (NC-stat)		0.80 (0.55-1.00)	0.55 (0.25-0.84)	1.8	0.37
Any tibial abnormality (NC-stat)		0.80 (0.55-1.00)	0.36 (0.08-0.65)	1.3	0.55
Tibial F-wave abnormality (NC-stat)		0.80 (0.55-1.00)	0.73 (0.46-0.99)	2.9	0.28
Tibial F-wave or CMAP abnormality (NC-stat)		0.80 (0.55-1.00)	0.64 (0.35-0.92)	2.2	0.31
Haig 2005 [33]	48				
Any needle examination abnormality		0.63 (0.43-0.82)	0.54 (0.34-0.74)	1.4	0.69
Any nerve conduction abnormality		0.54 (0.34-0.74)	0.75 (0.58-0.92)	2.2	0.61
Any abnormality		0.79 (0.63-0.95)	0.50 (0.30-0.70)	1.6	0.42
Paraspinal mapping >4		0.29 (0.11-0.47)	1.00	~	0.71
Haig 2006 [34]	82				
MRI		0.59	0.44 (0.27-0.61)	1.1	0.93
Electrodiagnosis		0.70 (0.57-0.83)	0.47 (0.30-0.64)	1.3	0.64
Haig 2007 [32]	82				
Abnormal MRI (minimum canal diameter ≤11.95 mm)		0.27 (0.15-0.40)	0.77 (0.63-0.92)	1.2	0.94

Electrodiagnosis: any abnormality		0.73 (0.60-0.85)	0.48 (0.31-0.66)	1.4	0.57
Fibrillations anywhere or absent H-wave		0.67 (0.54-0.80)	0.71 (0.55-0.87)	2.3	0.47
Fibrillations anywhere		0.53 (0.39-0.67)	0.87 (0.75-0.99)	4.1	0.54
Paraspinal fibrillations		0.33 (0.20-0.46)	0.90 (0.80-1.00)	3.4	0.74
Limb fibrillations		0.31 (0.19-0.44)	0.97 (0.91-1.00)	9.7	0.71
H-wave absent		0.63 (0.49-0.76)	0.29 (0.13-0.45)	0.88	1.3
Motor unit changes >2/10		0.41 (0.28-0.55)	0.61 (0.44-0.78)	1.1	0.96
Yagci 2009 [55]	60				
Paraspinal mapping >0		0.72 (0.55-0.88)	1.00	~	0.29
Paraspinal mapping >9		0.97 (0.90-1.00)	0.92 (0.81-1.00)	10.3	0.04
Paraspinal mapping >18		0.97 (0.90-1.00)	0.69 (0.53-0.85)	3.1	0.05
Paraspinal mapping > 27		0.97 (0.90-1.00)	0.63 (0.46-0.79)	2.6	0.06
Dermatomal Somatosensory Evoked Potentials (DSEP)					
Shen 2008 [52]	47	0.96	-	-	-
Snowden 1992 [53]	58				
Group without previous surgery	40	0.78 (0.64-0.88)	-	-	-
Group with previous surgery	18	0.70 (0.53-0.84)	-	-	-
Whole group	58	0.94 (0.71-1.00)	-	-	-
Magnetic stimulation MCT					
Han 2004 [35]	16*				
Total conduction time to abductor hallucis		0.22	-	-	-
Central motor conduction time to abductor hallucis		0.28	-	-	-
Caudal motor conduction time (MCT)		0.56	-	-	-
Tibial somatosensory evoked potential (SEP)l		0.44	-	-	-
Caudal MCT or SEP		0.66	-	-	-
Selective lumbar root sheath infiltration					
Castro 1991 [19]	30	**PPV 0.95 (0.74-1.00)**	-	-	-

Table 3 Diagnostic Accuracy of Clinical Tests for Lumbar Spinal Stenosis

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
		Age			
Cook 2010 [22]	1448				
Age >48		0.88 (0.85-0.89)	0.49 (0.47-0.50)	1.7	0.25
Katz 1995 [41]	75				
Age >65 years		0.77 (0.64-0.90)	0.69 (0.53-0.85)	2.5	0.33
Konno(1) 2007 [42]	468				
Age <60		0.15 (0.11-0.20)	0.62 (0.56-0.68)	0.41	1.4
Age >70		0.64 (0.58-0.71)	0.68 (0.62-0.74)	2.0	0.52
		Comorbidities			
Sugioka 2008 [57]	374				
Orthopaedic disease		0.18 (0.13-0.23)	0.91 (0.87-0.95)	2.0	0.90
		Pain locations			
Cook 2010 [22]	1448				
Bilateral symptoms		0.03 (0.02-0.04)	0.98 (0.98-0.99)	2.3	0.98
Leg pain more than back pain		0.16 (0.14-0.18)	0.92 (0.91-0.93)	2.1	0.91
Moderate back pain		0.95 (0.94-0.96)	0.02 (0.01-0.03)	1.0	2.1
Moderate buttock pain		0.81 (0.77-0.84)	0.33 (0.31-0.35)	1.2	0.60
Moderate leg pain		0.90 (0.88-0.92)	0.24 (0.22-0.25)	1.2	0.43
Pain constancy		0.23 (0.20-0.26)	0.78 (0.76-0.80)	1.1	0.98
Ljunggren 1991 [12]	179				
Pain relief with assuming a suitable body position		0.61 (0.50-0.72)	0.55 (0.45-0.65)	1.4	0.71
Bilateral buttock or leg pain		0.51 (0.40-0.62)	0.92 (0.87-0.97)	6.3	0.54
Lumbo-sacral pain		0.75 (0.65-0.84)	0.27 (0.18-0.36)	1.0	0.94
Gluteal pain		0.84 (0.75-0.92)	0.05 (0.01-0.09)	0.88	3.3
Thigh pain		0.95 (0.90-1.00)	0.14 (0.07-0.20)	1.1	0.37
Calf pain		0.91 (0.85-0.97)	0.06 (0.01-0.11)	0.97	1.5
Foot pain		0.57 (0.46-0.68)	0.27 (0.18-0.36)	0.78	1.6

Table 3 Diagnostic Accuracy of Clinical Tests for Lumbar Spinal Stenosis (continued)

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
Katz 1995 [41]	75				
Pain below buttocks		0.88 (0.78-0.98)	0.34 (0.18-0.50)	1.3	0.35
Pain below knees		0.56 (0.41-0.71)	0.63 (0.46-0.80)	1.5	0.70
Severe lower extremity pain		0.65 (0.51-0.79)	0.67 (0.51-0.83)	2.0	0.52
Roach 1997 [51]	99				
Radiating leg pain (disk disease with spinal stenosis)		0.94	0.21	1.2	0.29
Pseudoclaudication (spinal stenosis)		0.63	0.71	2.2	0.52
Pseudoclaudication (disc disease with spinal stenosis)		0.47	0.64	1.3	0.83
Symptoms reproduced by specific actions					
Cook 2010 [22]	1448				
Pain with walking/standing		0.67 (0.64-0.69)	0.44 (0.42-0.46)	1.2	0.75
Sitting relieves pain		0.26 (0.24-0.29)	0.86 (0.84-0.88)	1.9	0.86
Fritz 1997 [31]	45				
Pain in legs relieved by sitting		0.81 (0.61-0.93)	0.16 (0.03-0.40)	1.0	1.2
Better able to walk when holding a shopping cart		0.63 (0.38-0.84)	0.67 (0.35-0.90)	1.9	0.55
Sitting best posture with regard to symptoms		0.89 (0.70-0.98)	0.39 (0.17-0.64)	1.5	0.28
Walk/stand worst posture with regard to symptoms		0.89 (0.70-0.98)	0.33 (0.13-0.59)	1.3	0.33
Katz 1995 [41]	75				
No pain when seated		0.46 (0.30-0.62)	0.93 (0.84-1.00)	6.6	0.58
Symptoms improve when seated		0.52 (0.37-0.67)	0.83 (0.70-0.96)	3.1	0.58
Worse when walking		0.71 (0.57-0.85)	0.30 (0.14-0.46)	1.0	0.97
Konno(1) 2007 [42]	468				
Burning sensation around the buttocks and/or intermittent priapism associated with walking		0.06 (0.03-0.09)	0.99 (0.98-1.00)	7.2	0.95
Intermittent claudication		0.82 (0.77-0.87)	0.78 (0.73-0.83)	3.7	0.23
Exacerbation when standing up		0.68 (0.62-0.74)	0.70 (0.65-0.76)	2.3	0.45

Table 3 Diagnostic Accuracy of Clinical Tests for Lumbar Spinal Stenosis (continued)

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
Improvement when bending forward		0.52 (0.45-0.58)	0.92 (0.88-0.95)	6.4	0.52
Sugioka 2008 [57]	374				
Exacerbated while standing up		0.92 (0.88-0.96)	0.20 (0.14-0.27)	1.2	0.39
Other symptoms					
Konno(1) 2007 [42]	468				
Urinary disturbance		0.14 (0.09-0.19)	0.98 (0.96-1.00)	6.9	0.88
Numbness of perineal region		0.05 (0.02-0.07)	0.99 (0.97-1.00)	3.7	0.97
Bilateral plantar numbness		0.27 (0.21-0.33)	0.87 (0.83-0.92)	2.2	0.84
Sugioka 2008 [57]	374				
Treatment for symptoms needs to be repeated every year		0.40 (0.33-0.47)	0.81-0.75-0.86)	2.1	0.74
Wake up to urinate at night		0.86 (0.82-0.91)	0.27 (0.20-0.33)	1.2	0.51
Physical examination					
Cook 2010 [22]	1448				
Gait abnormality		0.29 (0.27-0.32)	0.81 (0.79-0.83)	1.6	0.87
Katz 1995 [41]	75				
Numbness		0.63 (0.49-0.74)	0.59 (0.42-0.76)	1.5	0.63
Poor balance		0.70 (0.56-0.84)	0.53 (0.36-0.70)	1.5	0.57
Wide-based gait		0.43 (0.28-0.58)	0.97 (0.91-1.00)	14	0.59
Abnormal Romberg		0.39 (0.24-0.54)	0.91 (0.81-1.00)	4.3	0.67
No pain with flexion		0.79 (0.67-0.91)	0.44 (0.27-0.61)	1.4	0.48
Thigh pain with 30 seconds of lumbar extension		0.51 (0.36-0.66)	0.69 (0.53-0.85)	1.6	0.71
Pinprick deficit		0.47 (0.32-0.62)	0.81 (0.67-0.95)	2.5	0.65
Weakness		0.47 (0.32-0.62)	0.78 (0.64-0.92)	2.1	0.68
Vibration deficit		0.53 (0.38-0.68)	0.81 (0.67-0.95)	2.8	0.58
Absent Achilles reflex		0.46 (0.31-0.61)	0.78 (0.64-0.92)	2.1	0.69
Konno(1) 2007 [42]	468				

Table 3 Diagnostic Accuracy of Clinical Tests for Lumbar Spinal Stenosis (continued)

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
Symptoms induced by having patients bend forward		0.18 (0.13-0.23)	0.63 (0.57-0.69)	0.47	1.3
Diagnostic support tools					
Cook 2010 [22]	1448				
1 of 5 positive findings (bilateral symptoms, leg pain more than back pain, pain during walking/standing, pain relief upon sitting, age >48)		0.96 (0.94-0.97)	0.20 (0.19-0.21)	1.2	0.19
2 of 5 positive findings		0.68 (0.65-0.71)	0.62 (0.60-0.64)	1.8	0.51
3 of 5 positive findings		0.29 (0.27-0.31)	0.88 (0.87-0.90)	2.5	0.80
4 of 5 positive findings		0.06 (0.05-0.07)	0.98 (0.98-0.99)	4.6	0.95
5 of 5 positive findings		<0.01 (0.001-0.003)	1.00 (0.99-1.00)	~	0.99
Kato 2009 [40]	118				
Clinical diagnostic support tool (LSS ≥7)		0.95 (0.89-1.00)	0.40 (0.28-0.52)	1.6	0.13
Konno(1) 2007 [42]	468				
Clinical diagnostic support tool (LSS ≥7)		0.93 (0.89-0.96)	0.72 (0.66-0.78)	3.3	0.10
Roach 1997 [51]	99				
Pain response to Activity and Position Questionnaire (spinal stenosis)		0.52	0.74	2.0	0.65
Pain response to Activity and Position Questionnaire (disk disease with spinal stenosis)		0.81	0.54	1.8	0.35
Sugioka 2008 [57]					
Derivation set (LSS ≥5)	374	0.81 (0.75-0.87)	0.58 (0.51-0.65)	1.9	0.33
Validation set (LSS ≥5)	94	0.75 (0.62-0.87)	0.51 (0.37-0.65)	1.5	0.50
Pain drawings					
Mann 1992 [45]	25	0.58 (0.41-0.73)	0.88 (0.83-0.93)	4.8	0.48
Mann 1991 [44]	250				

Table 3 Diagnostic Accuracy of Clinical Tests for Lumbar Spinal Stenosis (continued)

Source	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR	Negative LR
Computerized pain drawings categorized by a statistical analysis using traditional statistical algorithms					
Five category prediction		0.32 (0.29-0.35)	0.83 (0.82-0.84)	1.9	0.82
Two category prediction		0.34 (0.31-0.37)	0.83 (0.82-0.84)	2.0	0.80
Mann 1993 [46]	250				
Artificial Neural Network evaluation of computerized pain drawings					
Course input, modified intuitive source		0.42	-	-	-
Fine input, empirical source		0.48	-	-	-
Treadmill					
Fritz 1997 [31]	45				
Earlier onset of symptoms with level walking		0.68 (0.46-0.85)	0.83 (0.59-0.96)	4.0	0.39
Longer total walking time during inclined walking		0.50 (0.25-0.75)	0.92 (0.64-1.00)	6.3	0.54
Prolonged recovery after level walking		0.82 (0.60-0.95)	0.68 (0.43-0.87)	2.6	0.26
Model based on discriminant analysis in which the variables are time to onset of symptoms and recovery time		0.77 (0.56-0.91)	0.95 (0.74-1.00)	15	0.24
Jensen 1989 [38]	23				
Symptom march		0.63 (0.24-0.91)	0.80 (0.10-0.52)**?**	3.1	0.47
Bilateral paresis		0.38 (0.08-0.76)	0.87 (0.60-0.98)	2.9	0.72
Bilateral reflex changes		0.50 (0.16-0.84)	0.80 (0.52-0.96)	2.5	0.63
Any change of neurological status		1.00 (0.63-1.00)	0.33 (0.12-0.62)	1.5	0.00
Gait-analyses					
Papadakis 2009 [49]	70				
Cut-off 0.06 nats		0.97 (0.92-1.00)	0.80 (0.67-0.93)	4.9	0.04
Physiotherapist assessment					
Laslett 2005 [43]	13				
Physiotherapist assessment		0.23 (0.00-0.46)	-	-	-

Discussion

The purpose of the present study was to update a previously published systematic review on the diagnostic accuracy of tests used to diagnose LSS [8]; an additional 22 articles were included. Our updated review shows no superior accuracy for myelography compared to CT, MRI, or 3D-MRM, with MRI and 3D-MRM showing the highest sensitivity. The diagnostic accuracy of electrodiagnostic testing was only modest and showed no superior accuracy compared to MRI. Paraspinal mapping had a high specificity in two studies and it may have some utility in confirming the clinical significance of radiological LSS among subjects with atypical symptoms. Several clinical findings may be useful for the diagnosis of LSS, including radiating leg pain that is exacerbated while standing up, the absence of pain when seated, the improvement of symptoms when bending forward, and a wide-based gait. However, the accuracy of these findings has yet to be corroborated in properly designed confirmatory studies.

Quality

In the included studies, there was high heterogeneity in study design, diagnostic test of interest, test characteristics, patient characteristics, reference standard, and definition of lumbar spinal stenosis. Because of the heterogeneity of the studies, we refrained from statistical pooling.

The definition of LSS was often unclear or not specified at all (item 13). Furthermore, QUADAS items were frequently scored as unclear or inadequate because of poor reporting of data. In many older studies, specificity was not reported or could not be calculated. Without the corresponding specificity of a test one cannot make assumptions concerning the probability of having LSS. Shortcomings in design, data collection, and reporting affect the estimates of diagnostic accuracy, mostly resulting in an overestimation [59].

The recent studies more often had a prospective design, especially so for the studies of clinical tests. Furthermore, almost all recent studies avoided differential verification bias. Differential verification bias occurs when people with a positive index test receive another, often more invasive reference test, resulting in an overestimation of sensitivity and an underestimation of specificity.

Gold standard

Recent studies have shown us that there is a need for a consensus on criteria to define and classify lumbar spinal stenosis [5, 6]. A vague definition of an illness and imprecise criteria to either rule-in or rule-out an illness, poses a major problem when performing research in patients with such a disorder [6]. In the absence of widely accepted diagnostic criteria, almost all included studies devised their own construct. This limits the generalizability of findings. Further research on lumbar spinal stenosis is essential, but at a time when other musculoskeletal disease experts are considering revisions of well-established sets of criteria [60, 61], the absence of diagnostic and/or classification criteria in the field of lumbar spinal stenosis should be considered a major focus for international organizations and clinical investigators.

For our review, we regarded expert opinion based on clinical findings and imaging and/or surgery as the best available reference standard, according to current clinical practice. However, all imaging studies in this review used an anatomic standard, based on imaging and/or surgery findings. Surgical findings depend on positioning of the patient, and the clinical observation of the anatomy may be equivocal, depending on the examiner's views of how the clinical syndrome of LSS and its subtypes typically present. Besides, when surgical confirmation is used as a reference standard, blinding is usually infeasible, and verification bias is likely to be present [62]. It should also be noted that for those studies using imaging either as a reference standard or a diagnostic test, positioning may in theory also affect the appearance of stenosis, but the specific impact of postural dynamics on accuracy has not been well studied.

Recent studies about clinical tests used the consensus diagnosis of multiple expert spine clinicians as reference standard. However, this induces a problem with incorporation bias whereby the overall clinical findings are taken into account in establishing the diagnosis. Because a diagnosis of the clinical syndrome of LSS requires information from the clinical examination, such bias is unavoidable [11].

Limitations

Although a thorough search in Medline and Embase was performed, papers reporting on diagnostic tests of spinal stenosis different from those included in the review we present may have been missed. However, in the references of the included studies only one study not found with the

systematic search was identified. Therefore, it seems unlikely relevant diagnostic studies have been missed.

Conclusions for clinical practice

Further research on lumbar spinal stenosis is essential, but the absence of diagnostic and/or classification criteria should be considered a major focus for international organizations and clinical investigators. Furthermore, we recommend the use of a clinical reference standard.

Given the literature to date, at present the most promising imaging test for LSS is MRI, avoiding myelography because of its invasiveness and lack of superior accuracy. Electrodiagnostic studies showed no superior accuracy for conventional electrodiagnostic testing compared to MRI. These tests should be considered in the context of those presenting symptoms with the highest diagnostic value, including radiating leg pain that is exacerbated while standing up, the absence of pain when seated, the improvement of symptoms when bending forward, and a wide-based gait.

Supplemental digital content 1

MRI

Eight studies reported the sensitivity for MRI [13, 15, 17, 20, 38, 47, 49, 55] which ranged from 60% to 96%. Two recent studies reported the diagnostic accuracy of MRI to detect foraminal stenosis, the sensitivity ranged from 77% to 96%, the specificity ranged from 43% to 68% [13, 20]. One study reported the diagnostic accuracy of the nerve root sedimentation sign. A positive sedimentation sign was defined as the absence of sedimented lumbar nerve roots in the supine position. The results of this study suggested a sensitivity of 94% and a specificity of 100% [15]. In a study that used a three-sequence MRI protocol as a reference standard, a sensitivity of 60% was shown MRI with a single-sequence; the corresponding specificity was 95% [49].

3D-MR Myelography

Three studies reported the sensitivity for three-dimensional magnetic resonance myelography (3D-MRM) [13, 25, 30] which ranged from 60% to 100%. One study reported the diagnostic accuracy of 3D-MRM to detect foraminal stenosis, the sensitivity ranged from 60% to 96%, the specificity ranged from 84% to 99% [13].

CT

Seven studies reported the sensitivity for CT [14, 16, 18, 23, 27, 47, 55] which ranged from 21% to 100%. The specificity ranged from 60% to 98%.

Myelography

Ten studies reported the sensitivity for conventional myelography [14, 16-18, 25, 27, 30, 36, 38, 47], which ranged from 54% to 100%. Of the 9 studies that investigated myelography as well as CT, MRI, or 3D-MRM, 5 studies showed a higher sensitivity for MRI, 3D-MRM, or CT than for myelography [14, 17, 25, 27, 47], and 4 studies showed a similar sensitivity for myelography as for CT or MRI [14, 17, 18, 30]. A higher sensitivity for myelography was reported in 3 studies [16, 18, 38]. The specificity of myelography (88% and 91%) was slightly higher than that of CT and MRI (75%, 88% and 90%) in the two studies that reported the specificity of these tests [17, 47].
One study reported the sensitivity for CT-myelography, which was 87% [17]. One study reported the sensitivity for epidurography, which was 94% [55]. In addition, one study reported the sensitivity for

epidural venography, which was 77% [36].

Radiography

The sensitivity and specificity of plain radiography, as shown by one study in which CT was used as a reference standard, were 66% and 93%, respectively [24].

Ultrasound

Ultrasound of the lumbar spine was evaluated in two studies: one study used surgery as a reference standard, and the other study used myelography or CT as the reference standard [26, 53]. The sensitivity ranged from 90% to 95%, the specificity ranged from 96% to 100%.
One study investigated the diagnostic accuracy of ultrasound of the calcaneus to detect anatomic LSS [46]. The sensitivity was 52%, and the specificity was 70%.

Supplemental digital content 2

Electromyography and Nerve Conduction Studies

Three studies investigated the diagnostic accuracy of electromyography to detect LSS [21, 28, 29, 32-34, 54].
The first article by Haig et al. [33] evaluated the diagnostic accuracy of electromyography and nerve conduction studies using a clinical reference standard. The sensitivity of EMG was 63%; the specificity was 54%. The sensitivity of NCS was 54%; the specificity was 75%. The combined accuracy of EMG and NCS had a sensitivity of 79% and a specificity of 50%. In addition to the conventional EMG evaluation the article reported the diagnostic accuracy of paraspinal mapping; the sensitivity was 29%, the specificity was 100%.

The following articles by Haig et al. [32, 34] used clinical confirmation of LSS only as the reference standard without consideration of imaging and/or surgery findings. The overall evaluation of MRI had a sensitivity of 59% and a specificity of 44%. A minimum canal diameter of ≤ 11.95 mm had a sensitivity of 27% and a specificity of 77%. The sensitivity and specificity of EMG (including paraspinal mapping) and NCS combined were 73% and 48%, respectively. The findings on MRI were able to differentiate persons with LSS from asymptomatic subjects but not from persons with mechanical low-back pain, whereas electrodiagnosis was able to marginally discriminate all groups.
The article of Chiodo et al. assessed the asymptomatic subjects among the previously described study population [21]. The specificity for EMG was 59%; the specificity for MRI was 44%. There was no statistically significant relationship between the false positive rate of electrodiagnosis and MRI.
The study of Yagci et al. evaluated the diagnostic accuracy of paraspinal mapping using a clinical reference standard [54]. The sensitivity ranged from 72% to 97%, and the specificity ranged from 63% to 100% .

The study of Fisher et al. compared the diagnostic accuracy of conventional electrodiagnosis and computerized recording and analysis of EMG and NCS (NC-stat) [28, 29]. MRI or post-myelographic CT was used as a reference standard . The sensitivity for EMG was 60%, and the specificity was 82%. The sensitivity for EMG combined with NCS was 90%, and the specificity was 45%. The sensitivity for NC-stat

ranged from 60% to 90%, and the specificity ranged from 27% to 82%.

DSEP

Two studies investigated the diagnostic accuracy of dermatomal somatosensory-evoked potentials (DSEP). One study had CT or MRI as a reference standard [52]; this study showed a sensitivity of 96%, the specificity was not reported. The other study had surgery as a reference standard [51]; the sensitivity ranged from 78% to 94%, but the specificity was not reported.

Magnetic stimulation MCT

One study investigated the diagnostic accuracy of the caudal motor conduction time (caudal MCT) after magnetic stimulation. This study showed a sensitivity of 56% [35], the specificity was not reported.

Selective lumbar root sheath infiltration

One study investigated the diagnostic accuracy of selective lumbar root sheath infiltration with successful outcome of surgery as a reference standard. This study did not report a sensitivity or specificity but showed a positive predictive value of 95% [19].

Supplemental digital content 3

Age and Comorbidities

Three studies reported the sensitivity for the patient characteristic age [22, 40, 41]. The sensitivity of the patient characteristic younger than 60 years was 15%; the specificity was 62% [41]. The sensitivity of the patient characteristic older than 65 years was 77%; the specificity was 69% [40]. The sensitivity for the presence of orthopaedic conditions was 18%; the specificity was 91% [56].

Symptoms

Four studies reported the sensitivity for different pain locations [12, 22, 40, 50]. Radiating pain, calf pain, thigh pain, and moderate back pain all had a high sensitivity (>90%) [12, 22, 50]. Bilateral symptoms and leg pain that was worse than back pain had a specificity of 98% and 92% [12, 22]. Exacerbation while standing up had a sensitivity of 92% [56]; no pain when seated and improvement when bending forward had a specificity of 93% and 92% [40, 41].

The presence of symptoms related to cauda equina syndrome had a specificity of 99%. The specificity of urinary disturbance was 98% [41]. However, these symptoms were insensitive and present in only a small percentage of patients.

Physical examination

In general, physical examination tests had a lower sensitivity than clinical symptoms. One study found that the specificity of a wide-based gait was 97%, and the specificity of an abnormal Romberg test result was 91% [40].

Diagnostic support tools

Two studies used predictor variables that were independently associated with LSS to create risk scores for diagnosing LSS [39, 41, 56]. The sensitivity of a score of 7 or higher on a clinical diagnostic tool including history and examination findings was 95% [39] and 93% [41]; the specificity was 40% [39] and 72% [41]. Sensitivity was optimized by the combination of history and examination findings, but this resulted in a lower overall specificity. The sensitivity of a score of 5 or higher on a diagnostic tool including only questionnaire-based items was 81%; the specificity was 58% [56]. On testing in a validation sample, the sensitivity was 75%, and the specificity was 51% [56].

One study created a diagnostic tool to indicate the likelihood of

the presence of LSS, composed of patient characteristics and clinical symptoms [22]. Having 4 of 5 positive findings had a sensitivity of 6% and a specificity of 98%.
One other study investigated the diagnostic accuracy of a Pain Response to Activity and Position Questionnaire (PRAP) [50]. A positive PRAP had a sensitivity of 52%, and a specificity of 74%.

Pain drawings

One study investigated the diagnostic accuracy of pain drawings with a clinical reference standard [43-45]. The sensitivity of statistical analysis of the pain drawings in a two category prediction was 34%; the specificity was 83% [43]. The sensitivity of an expert evaluation of the pain drawings was 58%; the specificity was 88% [44].

Treadmill

Two studies investigated the diagnostic accuracy of a treadmill test. One study used CT or MRI as a reference standard; the sensitivity was 77%, and the specificity was 95% [31]. The other study had myelography as a reference standard; the sensitivity ranged from 38% to 100%, and the specificity ranged from 33% to 87% [37].

Gait-analyses

One study investigated the diagnostic accuracy of gait-analyses; this study showed a sensitivity of 97% and a specificity of 80% [48].

Physiotherapist assessment

One study investigated the diagnostic accuracy of a physiotherapist assessment to detect anatomic lumbar stenosis; this study showed a sensitivity of 23% [42].

Appendix Table 1 Study Characteristics of the Included Diagnostic Studies on Imaging Tests; 20 articles reporting 20 studies

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Aota 2007 [13]	117	117	63.4 (25-85)	65	Patients whose leg pain improved after selective decompression surgery. Normal volunteers as age- and gender- matched controls.	Setting: tertiary Design: retrospective	Foraminal stenosis	Surgery	MRI and MR-myelography
Arrault 1987 [14]	60	60	Unknown	58	Surgically confirmed central and lateral spinal stenosis; symptoms: unknown	Setting: tertiary Design: retrospective	Central stenosis, lateral stenosis (only osseous or associated with herniated disc)	Surgery	CT and myelography
Barz 2010 [15]	200	200	(49-74)	47	Consecutive patient with symptoms of nonspecific LBP, leg pain or claudication. Patients with LSS at level L5/S1 were excluded	Setting: secondary and tertiary Design: retrospective	Central canal stenosis	MRI: CSA of the dural sac	MRI: nerve root sedimentation sign
Bell 1984 [16]	122	46	Unknown	Unknown	Surgically confirmed disc herniation (n=76), spinal stenosis (n=46) or both; symptoms: unknown	Setting: unknown Design: retrospective	Spinal stenosis including nerve compression due to facet joint abnormalities and lateral recess stenosis	Surgery	CT and myelography
Bischoff 1993 [17]	57	28	(20-79)*	51*	Surgically explored for suspected disc herniation (n=47) or spinal stenosis (n=28); symptoms: unknown	Setting: tertiary Design: retrospective	Spinal stenosis including nerve root compression due to facet joint arthritis, foraminal, lateral recess stenosis as one group	Surgery	CT-myelography, MRI and myelography

Appendix Table 1 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Bolender 1985 [28]	55	24	(36-85)*	Unknown	Surgery for suspected central lateral and foraminal stenosis; symptoms: pain, neurogenic claudication, sensory changes, weakness and absence of reflexes in the lower extremities	Setting: tertiary Design: retrospective	Central stenosis	Surgery	CT and myelography
Chang 2010 [20]	39	13	69.2*	33.3*	Patients who underwent surgery for lumbar foraminal stenosis; symptoms: in most cases unilateral pain	Setting: tertiary Design: retrospective	Foraminal stenosis	Excellent or good result (MacNab) of surgery	Coronal thin-sliced MRI
Donmez 1990 [23]	50	50	(21-61)	56	Patients were preoperatively evaluated by CT; symptoms: low back pain, leg pain, paresthesia, neurogenic claudication	Setting: unknown Design: prospective	Lateral recess and central spinal stenosis	Surgery	CT
Eberhardt 1994 [25]	65	65	51.8 (23-80)	42	Surgery for lumbar complaints; symptoms: radicular complaints	Setting: unknown Design: prospective	Spinal stenosis: osseous narrowing of the spinal canal, with liquor disturbance, subgroup is spondylolisthesis	Surgery	Myelography and 3D-MR myelography
Engel 1985 [26]	67	19	Unknown	Unknown	Symptomatic spine patients, not all patients had symptomatology warranting myelogram or surgery; symptoms: low back pain, with or without radiation down one or both legs	Setting: tertiary Design: unknown	Focal stenosis isolated or superimposed on diffuse (including herniated disc)	Surgery	Ultrasound

Appendix Table 1 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Feldmeyer 1982 [27]	38	7	(16-81)*	55*	Suspected disc herniation (n=18), or canal stenosis (n=13), suspected for other spinal diseases (n=7) with indication of preoperative myelography; symptoms: bilateral radiation, sciatica, neurogenic claudication and absence of Lasegue	Setting: unknown Design: prospective	Canal stenosis including narrow canal only, or associated with herniated disc or bony protrusions	Surgery	CT and myelography
Freund 1997 [30]	25	25	44	64	Clinical indication for myelography, suspected for spinal canal stenosis; symptoms: unknown	Setting: unknown Design: prospective	Degenerative osseous spinal canal stenosis	Surgery	Myelography and 3D-MR myelography
Herkowitz 1982 [36]	30	18	59 (21-95)	50	Surgically confirmed disc herniation (n=12) or stenosis (n=18); symptoms: sciatica	Setting: unknown Design: retrospective	Spinal stenosis	Surgery	Myelography and epidural venography
Ilkko 1988 [24]	116	116	44 (10-79)	58	Indication for radiologic examination, because of: sciatica (n=69), suspected spinal stenosis (n=18), chronic lumbar pain (n=15), spondylosis and spondylolisthesis (n=4), suspected fracture, pain associated with degeneration, piriformis syndrome, discitis, polyneuropathy and MS (n=10)	Setting: tertiary Design: prospective	Central spinal stenosis, including bony and soft tissue (disc protrusion, ligamentum flavum)	CT	Radiography

Appendix Table 1 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Jia 1991 [38]	78	27	41 (25-67)*	65*	Surgically confirmed disc herniation (n=65), including 10 with concurrent osseous lateral recess stenosis and 4 with hypertrophic ligamentum flavum; nerve root canal or lateral recess stenosis only (n=8) and central stenosis (n=5); symptoms: unknown	Setting: unknown Design: unknown	Central canal, nerve root canal, lateral recess stenosis as one group	Surgery	Myelography and MRI
Mariconda 2004 [46]	117	117	60.1 (40-70+)	44.4	Hospitalised or attending an outpatient clinic for low back pain; symptoms: low back pain either radiating down the leg or not	Setting: secondary and tertiary Design: cross-sectional	Central canal stenosis	MRI (T1 sagittal, T2 transverse, T2 sagittal) and anteroposterior and lateral radiography	Ultrasound calcaneus
Modic 1986 [47]	60	48	46 (19-73)	Unknown	Clinical history and physical exam that indicated a strong probability of disc herniation or canal stenosis with a likelihood of required surgery; symptoms: unknown	Setting: unknown Design: prospective	Neural foramina, lateral recess, central canal stenosis as one group	Surgery	MRI, CT and myelography
Rankine 1997 [49]	79	79	48.8	42	Indication for MRI; symptoms: low back pain, with or without sciatica	Setting: unknown Design: retrospective	Central canal stenosis	MRI: T1 sagittal, T2 axial, T2 sagittal	MRI

Appendix Table 1 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Tervonen 1989 [53]	76	76	42 (21-63)	45	Indication for examination by CT or myelography; symptoms: unknown	Setting: tertiary Design: prospective	Central spinal stenosis	Myelography or CT	Ultrasound
Yan 2010 [55]	29	22	57.3 (36-71)*	59*	Inclusion criteria: central canal or nerve root canal stenosis confirmed by CT and MRI scan with clinical symptoms, the lateral herniated nucleus pulposus confirmed by CT or MRI scans, with irritation sign of the nerve root, typical clinical symptoms of lumbar spinal stenosis but no powerful findings on CT or MRI scans, the postoperative recurrence of a lumbar decompression procedure	Setting: tertiary Design: prospective	Nerve root compression due to a combination of disk herniation, redundancy and hypertrophy of the ligamentum flavum and hypertrophy of the facet joints with accompanying osteophytes	Surgery	MRI, CT and multispiral CT epidurography

* Reported results based on the total group

Appendix Table 2 Study Characteristics of the Included Diagnostic Studies on Electrodiagnostic Tests; 11 articles reporting 7 studies

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Castro 1991 [19]	30	30	53	Unknown	Suspicion of nerve root compression by degenerative changes; symptoms: upper leg pain, aggravating when standing, typical neurogenic claudication (n=9), segmental pain in a dermatome (n=12), long history of back complaints	Setting: unknown Design: prospective	Degenerative stenosis of the nerve root canal or lateral recess, as one group	Successful outcome of surgery	Selective lumbar root sheath infiltration
Fisher 2007 2008 [28,29]	34	21	59.8 (39-89)*	97.1*	Consecutive patients referred for evaluation of a possible lumbosacral radiculopathy; symptoms: low back pain or buttock pain that was associated with pain, numbness, and/or paresthesias of one or both lower extremities and/or consistent neurological findings on examination	Setting: secondary Design: retrospective	Foraminal stenosis, central canal stenosis	MRI and post-myelographic CT	Electrodiagnostic testing including NC-stat

Appendix Table 2 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Haig 2005 Haig 2006 Haig 2007 Chiodo 2007 [21,32, 33,34]	150	**A** 48 **B** 126 **C** 126 **D** 32	(55-80)*	Unknown	Participants with varying severity of spinal stenosis based on radiologist reports, low back pain patients with no MRI evidence of spinal stenosis and asymptomatic volunteers. **A** 'Gold standard' population in whom there was complete agreement regarding the diagnosis **B, C** Participants for whom complete data was available **D** Asymptomatic volunteers	Setting: secondary Design: prospective	Lumbar spinal stenosis	**A** Expert unanimous agreement by clinical and radiological data (MRI) **B, C, D** Clinical diagnoses by physiatrist's impression of spinal history and physical examination	**A** Electrodiagnostic testing **B, C, D** Electrodiagnostic testing and MRI
Han 2004 [35]	28	16	56.3	Unknown	Confirmed narrowing of the spinal canal or neural foramina by MRI; symptoms: pain in the buttocks or legs that was aggravated by walking and relieved after sitting, weakness and tingling of the lower legs	Setting: secondary Design: cross-sectional	Spinal canal stenosis and neural foramina stenosis	MRI	Caudal MCT (after magnetic stimulation)

Appendix Table 2 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Shen 2008 [51]	97	47	53 (42-73)	30	Diagnosed with lumbar spinal stenosis by CT/ MRI and confirmed by surgery; symptoms: long-term lumbago combined with radiating pain in both lower extremities and intermittent claudication	Setting: tertiary Design: retrospective	Lumbar spinal stenosis	Surgery	Dermatomal somatosensory evoked potential (DSEP)
Snowden 1992 [52]	58	A 40 B 18 C 58	64 (23-85)	69	High index of suspicion for spinal stenosis by history and physical examination, indication for DSEP, previously examined by MRI and/or CT; symptoms: unknown	Setting: tertiary Design: retrospective	Central, lateral, recess or foraminal stenosis, as one group	CT and/or MRI	Dermatomal somatosensory evoked potential (DSEP)
Yagci 2009 [54]	62	60	55.4 (40-70)	Unknown	CR-LSS group with MRI confirmed lumbar spinal stenosis with mechanical low-back pain, buttock pain, or neurogenic claudication >6 months. R-LSS group with MRI confirmed spinal stenosis and no low-back pain or neurogenic claudication. Control group with low-back pain but not clinical symptoms or radiological findings of central or lateral recess stenosis	Setting: secondary Design: prospective	Lumbar spinal stenosis	Clinical assessment by clinical and radiological data	Paraspinal mapping

* Reported results based on the total group

Appendix Table 3 Study Characteristics of the Included Diagnostic Studies on Clinical Tests; 15 articles reporting 12 studies

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Cook 2010 [22]	1448	1448	55.3	40.5	Consecutive patients with suspicion of a condition associated with origin at the lumbar spine	Setting: tertiary Design: prospective	Central, lateral recess and foraminal stenosis	Expert opinion based on clinical findings and imaging (MRI)	Standardized history and physical examination
Fritz 1997 [31]	45	45	58	Unknown	All patients had previously undergone MRI or CT; symptoms: low back and lower extremity pain and selfreported limitations in walking tolerance	Setting: tertiary Design: prospective	Spinal stenosis: any narrowing of the spinal canal and/or nerve root canals	MRI or CT	1 Two stage treadmill test 2 Medical history questionnaire
Jensen 1989 [37]	23	23	55 (23-72)	57	Indication for myelography; symptoms: neurogenic claudication, unilateral in 14 cases, bilateral in 9 cases	Setting: unknown Design: prospective	Spinal stenosis	Myelography	Treadmill, downhill walking
Kato 2009 [39]	119	118	68.2 (12-96)	52.5	Patients with symptoms in lower extremities Inclusion of non-adults. We were unable to collect the data for adults only from the author	Setting: secondary Design: prospective	Narrowing of the canal with the posterior element as the main factor	Consensus diagnostic impression of expert physicians, confirmation by x-rays, CT and MRI	Clinical diagnostic tool including history and physical examination

Appendix Table 3 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Katz 1995 [40]	93	75	65 (40-91)	31	Patients with low back pain with or without radiation to the lower extremities	Setting: tertiary Design: prospective	Spinal stenosis: compression of nerve roots by narrowing of the spinal canal or neural foramina	Expert opinion (>80% confidence in diagnosis for cases and <20% for non-cases)	Standardized history and physical examination
Konno(1) 2007 Sugioka 2008 [41,56]	469	**A** 468 **B** 374	65.2 (20-96)*	54.2*	Consecutive patients showing primary symptoms of pain or numbness in the lower extremities, including the buttocks, thighs and lower legs	Setting: primary and secondary care Design: prospective	Lumbar spinal stenosis	Consensus diagnostic impression of expert physicians, confirmation by x-rays and MRI	**A** History and physical examination **B** Questionnaire items
Konno(2) 2007 [57]	250	250	63.2	48.4	Consecutive patients with primary symptoms of pain or numbness in the legs This study is presented very poorly, many aspects are unclear or contradictory	Setting: primary and secondary care Design: prospective	Lumbar spinal stenosis	Consensus diagnostic impression of expert physicians, confirmation by x-rays and MRI	Questionnaire items

Appendix Table 3 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Laslett 2005 [42]	216	13	44.2 (20-77)*	56.9*	Consecutive patients with chronic lumbopelvic pain and/or referred lower extremity symptoms scheduled to receive standards examinations	Setting: secondary Design: prospective	Spinal stenosis	Fluoroscopically guided injections and imaging assessment by clinical expert	Comprehensive physiotherapist assessment
Ljunggren 1991 [12]	179	179	Unknown	58.1	Consecutive patients with lumbago-sciatica and no previous back surgery. They were considered potential cases for lumbar surgery	Setting: secondary Design: unknown	Canal and recess stenosis	Diagnosis by physical and neurological examination, imaging and surgery	McGill Pain Questionnaire
Mann 1991 1992 1993 [43,44,45]	250	**A** 25 **B+C** 250	Unknown	Unknown	Five categories of patients were selected, with 50 patients in each category: benign back pain, herniated nucleus pulposus, spinal stenosis, serious underlying disorders or psychogenic regional pain disturbance; symptoms: unknown	Setting: tertiary Design: retrospective	Spinal stenosis: narrowing of the spinal canal and foramina of congenital, developmental or degenerative origin	Review of clinical notes, image findings and surgical reports, if applicable. Verification of final diagnosis through the course of treatment	Computerized pain drawings

Appendix Table 3 (continued)

	N	Participants on Whom Results are Based (N)	Mean Age (range)	% Male	Characteristics of Participants	Setting and Design	Type stenosis	Reference test	Index test
Papadakis 2009 [48]	70	70	50.9	48.6	The experimental group's subjects were diagnosed with spinal stenosis, without any other neuromuscular and musculoskeletal pathology or injury, using MRI scans. The control group consisted of healthy subjects with no history of neuromuscular and musculoskeletal pathology or injury.	Setting: unknown Design: case-control (prospective)	Central canal stenosis	Anterior-posterior diameter of spinal canal on MRI	Gait variability using a tri-axial accelerometer
Roach 1997 [50]	106	99	55 (19-88)	51	Symptoms: recurrent or chronic low back pain	Setting: tertiary Design: prospective	Spinal stenosis: narrowing of the spinal canal or foramina of developmental or degenerative origin	Combination of patients complaints, results of physical examination, radiologic tests, laboratory work, consultations and other tests	1 Pain response to Activity and Position Questionnaire 2 radiating leg pain 3 pseudoclaudication

* Reported results based on the total group

Item 1*	Was the spectrum of patients representative of the patients who will receive the test in practice?
Item 2*	Is the reference standard likely to classify the target condition correctly?
Item 3*	Is the time period between the reference standard and index test short enough to be reasonably sure that the target condition did not change in the time between the two tests?
Item 4	Did the whole sample or a random selection of the sample, receive verification using the intended reference standard?
Item 5	Did patients receive the same reference standard irrespective of the index test result?
Item 6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
Item 7	Were the reference standard results interpreted without knowledge of the results of the index test?
Item 8	Were the index test results interpreted without knowledge of the results of the reference standard?
Item 9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
Item 10*	Were uninterpretable/ intermediate test results reported for both reference test and index test?
Item 11	Were withdrawals from the study explained?
Additional QUADAS	
Item 12	Was the index test applied correctly? The execution of the index test should be described in sufficient detail to permit replication of the test
Item 13*	Were cut-off values/gradings or "positive results" of the reference test clearly defined?
Item 14	Were cut-off values/gradings or "positive results" of the index test clearly defined?
Item 15*	Were data on interobserver variation reported and within an acceptable range?

Appendix Figure 1. The items of the QUADAS tool for methodological assessment of diagnostic studies

*Item 1: All people should have symptoms of lumbar spinal stenosis. People should not have confirmed lumbar spinal stenosis at the start of the study.
*Item 2: Surgical confirmation or a combination of clinical diagnosis (low back pain and claudicatio) and radiologic imaging methods together were used as a reference standard.
*Item 3: The time period is less than one year.
*Item 10: Uninterpretable of both reference test and index test.
*Item 13: Studies with a reference test of expert opinion/consensus are scored not applicable (NA)
*Item 15: Acceptable range: ICC >0.75, kappa >0.60

	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?	Index test described in sufficient detail?	Cut-off value of reference test defined?	Cut-off value of index test defined?	Interobserver variation reported?
Imaging															
Aota 2007 [15]	-	+	?	-	+	+	+	+	-	+	+	+	-	+	+
Arrault 1987 [16]	-	+	?	-	-	+	-	+	-	-	?	-	-	-	?
Barz 2010 [17]	-	+	+	+	+	-	+	-	+	+	+	+	+	+	+
Bell 1984 [18]	-	+	?	-	-	+	-	+	-	-	?	-	-	-	?
Bischoff 1993 [19]	?	+	?	-	-	+	-	+	-	-	?	-	+	+	?
Bolender 1985 [20]	-	+	?	-	-	+	-	?	?	-	?	+	?	+	?
Chang 2010 [22]	-	?	?	+	+	+	?	?	?	?	?	+	+	+	?
Donmez 1990 [25]	+	+	?	-	-	+	-	+	?	-	?	-	-	-	?
Eberhardt 1994 [27]	+	+	+	+	+	+	-	?	?	-	-	+	-	-	?
Engel 1985 [28]	+	+	?	-	-	+	?	+	+	-	-	+	-	-	?
Feldmeyer 1982 [29]	+	+	?	-	-	+	-	+	?	-	-	-	-	-	?
Freund 1997 [32]	+	+	?	?	?	+	-	+	-	?	-	+	-	-	?
Herkowitz 1982 [38]	-	+	?	-	-	+	-	?	?	-	?	-	-	-	?
Ilkko 1989 [39]	-	-	+	+	+	+	+	+	?	-	-	+	+	+	?
Jia 1991 [41]	-	+	?	-	-	+	-	?	?	-	?	-	-	-	?
Mariconda 2004 [49]	-	-	?	+	+	+	+	?	+	+	+	+	+	+	?
Modic 1986 [50]	+	+	?	-	-	+	-	+	-	-	-	+	-	-	?
Rankine 1997 [52]	+	-	+	+	+	-	+	+	-	-	-	+	+	+	-
Tervonen 1989 [56]	+	-	+	+	-	+	+	+	+	-	?	+	+	+	?
Yan 2010 [58]	-	+	?	-	?	?	+	?	?	?	+	?	+	+	?
Electrodiagnostic															
Castro 1991 [21]	?	-	?	-	-	+	-	+	+	-	+	+	+	+	?
Fisher 2007, 2008 [30,31]	+	+	+	-	+	+	+	?	+	?	-	-	-	-	NA
Haig 2005,2006,2007 Chiodo 2007 [23,34,35,36]	+	+	?	+	+	+	+	+	-	+	+	+	NA	+	NA
Han 2004 [37]	-	+	?	-	+	+	+	?	?	?	+	+	-	+	?
Shen 2008 [54]	-	+	?	-	+	+	?	-	+	?	+	-	-	+	?
Snowden 1992 [55]	+	-	?	-	-	+	+	+	?	-	?	+	?	+	?
Yagci 2009 [57]	-	+	?	+	+	+	+	+	-	+	+	-	NA	-	NA
Clinical															
Cook 2010 [24]	-	-	?	+	?	-	-	?	+	?	?	+	NA	+	?
Fritz 1997 [33]	+	-	?	+	-	+	+	?	?	-	-	+	-	+	?
Jensen 1989 [40]	+	-	?	+	+	+	+	+	+	-	?	-	+	+	?
Kato 2009 [42]	+	-	?	+	+	-	-	+	+	+	+	+	NA	+	?
Katz 1995 [43]	-	-	?	+	+	-	+	+	+	?	-	-	NA	+	?
Konno(1) 2007 Sugioka 2008 [44,59]	+	+	?	+	+	+	+	+	+	+	+	+	NA	+	-
Konno(2) 2007 [60]	-	-	?	-	+	+	+	-	?	?	-	+	NA	-	-
Laslett 2005 [45]	-	-	+	-	+	+	+	+	+	?	+	?	-	+	-
Ljunggren 1991 [14]	+	+	?	+	?	+	?	+	+	?	?	+	NA	+	?
Mann 1991,1992,1993 [46,47,48]	-	+	?	+	-	+	+	+	+	-	-	?	NA	?	-
Papadakis 2009 [51]	-	?	?	?	+	+	?	?	?	?	?	+	-	+	?
Roach 1997 [53]	+	+	?	+	-	-	+	+	+	-	+	?	NA	?	+

Appendix Figure 2. Review author's judgements about each methodological quality item for each included study
? = Unclear
NA = Not applicable

References

1. Genevay, S. and S.J. Atlas, Lumbar spinal stenosis. *Best Pract Res Clin Rheumatol*, 2010. 24(2): p. 253-65.
2. Boden, S.D., et al., Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am*, 1990. 72(3): p. 403-8.
3. Kalichman, L., et al., Spinal stenosis prevalence and association with symptoms: the Framingham Study. *Spine J*, 2009. 9(7): p. 545-50.
4. Wiesel, S.W., et al., A study of computer-assisted tomography. I. The incidence of positive CAT scans in an asymptomatic group of patients. *Spine (Phila Pa 1976)*, 1984. 9(6): p. 549-51.
5. Genevay, S., S.J. Atlas, and J.N. Katz, Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: a structured literature review. *Spine (Phila Pa 1976)*, 2010. 35(7): p. 803-11.
6. Steurer, J., et al., Quantitative radiologic criteria for the diagnosis of lumbar spinal stenosis: a systematic literature review. *BMC Musculoskelet Disord*, 2011. 12(1): p. 175.
7. Kent, D.L., et al., Diagnosis of lumbar spinal stenosis in adults: a metaanalysis of the accuracy of CT, MR, and myelography. *AJR Am J Roentgenol*, 1992. 158(5): p. 1135-44.
8. de Graaf, I., et al., Diagnosis of lumbar spinal stenosis: a systematic review of the accuracy of diagnostic tests. *Spine (Phila Pa 1976)*, 2006. 31(10): p. 1168-76.
9. Whiting, P., et al., Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess*, 2004. 8(25): p. iii, 1-234.
10. Juni, P., et al., The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*, 1999. 282(11): p. 1054-60.
11. Suri, P., et al., Does this older adult with lower extremity pain have the clinical syndrome of lumbar spinal stenosis? *JAMA*, 2010. 304(23): p. 2628-36.
12. Ljunggren, A., Discriminant validity of pain modalities and other sensory phenomena in patients with lumbar herniated intervertebral discs versus lumbar spinal stenosis. *Neuro-Orthopedics*, 1991(11): p. 91-99.
13. Aota, Y., et al., Magnetic resonance imaging and magnetic resonance myelography in the presurgical diagnosis of lumbar foraminal stenosis. *Spine (Phila Pa 1976)*, 2007. 32(8): p. 896-903.

14. Arrault, I., et al., [Diagnostic value of saccoradiculography and scanner in lumbar stenoses] Valeur diagnostique de la saccoradiculographie et du scanner dans les stenoses lombaires. *Rev Rhum Mal Osteoartic*, 1987. 54(10): p. 627-31.
15. Barz, T., et al., Nerve root sedimentation sign: evaluation of a new radiological sign in lumbar spinal stenosis. *Spine (Phila Pa 1976)*, 2010. 35(8): p. 892-7.
16. Bell, G.R., et al., A study of computer-assisted tomography. II. Comparison of metrizamide myelography and computed tomography in the diagnosis of herniated lumbar disc and spinal stenosis. *Spine (Phila Pa 1976)*, 1984. 9(6): p. 552-6.
17. Bischoff, R.J., et al., A comparison of computed tomography-myelography, magnetic resonance imaging, and myelography in the diagnosis of herniated nucleus pulposus and spinal stenosis. *J Spinal Disord*, 1993. 6(4): p. 289-95.
18. Bolender, N.F., N.S. Schonstrom, and D.M. Spengler, Role of computed tomography and myelography in the diagnosis of central spinal stenosis. *J Bone Joint Surg Am*, 1985. 67(2): p. 240-6.
19. Castro, W.H. and P.F. van Akkerveeken, [The diagnostic value of selective lumbar nerve root block] Der diagnostische Wert der selektiven lumbalen Nervenwurzelblockade. *Z Orthop Ihre Grenzgeb*, 1991. 129(4): p. 374-9.
20. Chang, H.S., et al., Microsurgical Posterolateral Transmuscular Approach for Lumbar Foraminal Stenosis. *J Spinal Disord Tech*, 2010.
21. Chiodo, A., et al., Needle EMG has a lower false positive rate than MRI in asymptomatic older adults being evaluated for lumbar spinal stenosis. *Clin Neurophysiol*, 2007. 118(4): p. 751-6.
22. Cook, C., et al., The clinical value of a cluster of patient history and observational findings as a diagnostic support tool for lumbar spine stenosis. *Physiother Res Int*, 2010.
23. Donmez, T., et al., Diagnostic value of computed tomography in spinal and lateral recess stenosis, preoperatively and for long-term follow-up: a prospective study in 50 cases. *Radiat Med*, 1990. 8(4): p. 111-5.
24. E., I., Diagnosis of lumbar central spinal stenosis by plain radiography. *J Med Imag*, 1989(3): p. 91-101.
25. Eberhardt, K.E., H.P. Hollenbach, and W.J. Huk, [3D-MR myelography in diagnosis of lumbar spinal nerve root compression syndromes. Comparative study with conventional myelography] 3D-MR Myelographie (3D-MRM) zur Diagnose von lumbalen Nervenwurzelkompressionssyndromen. Vergleichsstudie zu konventioneller Myelographie. *Aktuelle Radiol*, 1994. 4(6): p. 313-7.

26. Engel, J.M., G.M. Engel, and D.R. Gunn, Ultrasound of the spine in focal stenosis and disc disease. *Spine (Phila Pa 1976)*, 1985. 10(10): p. 928-31.
27. Feldmeyer JJ, S.P., Regli F, et al., Apport de la tomodensitometrie au diagnostic de la hernie discale et du canal etroit lombaires. *Schweiz Med Wochenshr*, 1982(112): p. 599-604.
28. Fisher, M.A., R. Bajwa, and K.N. Somashekar, Lumbosarcral radiculopathties--the importance of EDX information other than needle electromyography. *Electromyogr Clin Neurophysiol*, 2007. 47(7-8): p. 377-84.
29. Fisher, M.A., R. Bajwa, and K.N. Somashekar, Routine electrodiagnosis and a multiparameter technique in lumbosacral radiculopathies. *Acta Neurol Scand*, 2008. 118(2): p. 99-105.
30. Freund, M., et al., [MR myelography in spinal canal stenosis] MR-Myelographie bei Spinalkanalstenosen. *Rofo*, 1997. 167(5): p. 474-8.
31. Fritz, J.M., et al., Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. *J Spinal Disord*, 1997. 10(5): p. 410-6.
32. Haig, A.J., et al., Electromyographic and magnetic resonance imaging to predict lumbar stenosis, low-back pain, and no back symptoms. *J Bone Joint Surg Am*, 2007. 89(2): p. 358-66.
33. Haig, A.J., et al., The sensitivity and specificity of electrodiagnostic testing for the clinical syndrome of lumbar spinal stenosis. *Spine (Phila Pa 1976)*, 2005. 30(23): p. 2667-76.
34. Haig, A.J., et al., Spinal stenosis, back pain, or no symptoms at all? A masked study comparing radiologic and electrodiagnostic diagnoses to the clinical impression. *Arch Phys Med Rehabil*, 2006. 87(7): p. 897-903.
35. Han, T.R., et al., A new method to measure caudal motor conduction time using magnetic stimulation. *Muscle Nerve*, 2004. 30(6): p. 727-31.
36. Herkowitz, H.N., et al., Metrizamide myelography and epidural venography. Their role in the diagnosis of lumbar disc herniation and spinal stenosis. *Spine (Phila Pa 1976)*, 1982. 7(1): p. 55-64.
37. Ilkko, Diagnosis of lumbar central spinal stenosis by plain radiography. *J Med Imag*, 1989(3): p. 91-101.
38. Jensen OH, S.-O.S., A new functional test in the diagnostic evaluation of neurogenic intermittent claudication. *Clin Rheumatol*, 1989(8): p. 363-367.
39. Jia, L.S. and Z.R. Shi, MRI and myelography in the diagnosis of lumbar canal stenosis and disc herniation. A comparative study. *Chin Med J (Engl)*, 1991. 104(4): p. 303-6.
40. Kato, Y., et al., Validation study of a clinical diagnosis support tool for lumbar spinal stenosis. *J Orthop Sci*, 2009. 14(6): p. 711-8.

41. Katz, J.N., et al., Degenerative lumbar spinal stenosis. Diagnostic value of the history and physical examination. *Arthritis Rheum*, 1995. 38(9): p. 1236-41.
42. Konno, S., et al., Development of a clinical diagnosis support tool to identify patients with lumbar spinal stenosis. *Eur Spine J*, 2007. 16(11): p. 1951-7.
43. Laslett, M., et al., Agreement between diagnoses reached by clinical examination and available reference standards: a prospective study of 216 patients with lumbopelvic pain. *BMC Musculoskelet Disord*, 2005. 6: p. 28.
44. Mann NH 3rd, B.M., Enger I, Statistical diagnosis of lumbar spine disorder using computerized patient pain drawings. *Comput Biol Med*, 1991(21): p. 393-397.
45. Mann NH 3rd, B.M., Enger I, Expert performance in low-back disorder recognition using patient drawings. *J Spinal Disord*, 1992(5): p. 254-259.
46. Mann NH 3rd, B.M., Hertz DB, Initial-impression diagnosis using low-back pain patient pain drawings. *Spine*, 1993(18): p. 41-53.
47. Mariconda, M., et al., Quantitative ultrasound measurements of the calcaneus in the prediction of lumbar spine degeneration. *Eur Spine J*, 2004. 13(4): p. 346-353.
48. Modic, M.T., et al., Lumbar herniated disk disease and canal stenosis: prospective evaluation by surface coil MR, CT, and myelography. *AJR Am J Roentgenol*, 1986. 147(4): p. 757-65.
49. Papadakis, N.C., et al., Gait variability measurements in lumbar spinal stenosis patients: part A. Comparison with healthy subjects. *Physiol Meas*, 2009. 30(11): p. 1171-86.
50. Rankine, J.J., C.E. Hutchinson, and D.G. Hughes, MRI of lumbar spondylosis: a comparison of sagittal T2 weighted and three sequence examinations. *Br J Radiol*, 1997. 70(839): p. 1112-21.
51. Roach KE, B.M., Albin RD, The sensitivity and specificity of pain response to activity and position in categorizing patients with low back pain. *Phys Ther*, 1997(77): p. 730-738.
52. Shen, N., et al., Evaluation of degree of nerve root injury by dermatomal somatosensory evoked potential following lumbar spinal stenosis. *Neural Regen Res*, 2008. 3(11): p. 1249-1252.
53. Snowden, M.L., et al., Dermatomal somatosensory evoked potentials in the diagnosis of lumbosacral spinal stenosis: comparison with imaging studies. *Muscle Nerve*, 1992. 15(9): p. 1036-44.
54. Tervonen, O. and J. Koivukangas, Transabdominal ultrasound measurement of the lumbar spinal canal. Its value for evaluation of lumbar spinal stenosis. *Spine (Phila Pa 1976)*, 1989. 14(2): p. 232-5.

55. Yagci, I., et al., The utility of lumbar paraspinal mapping in the diagnosis of lumbar spinal stenosis. *Am J Phys Med Rehabil*, 2009. 88(10): p. 843-51.
56. Yan, L., et al., The study of epidurography and multispiral CT scanning examinations in the diagnosis of lumbar nerve root canal stenosis. *Orthopedics*, 2010. 33(10): p. 732.
57. Sugioka, T., et al., Predictive value of self-reported patient information for the identification of lumbar spinal stenosis. *Fam Pract*, 2008. 25(4): p. 237-44.
58. Konno, S., et al., A diagnostic support tool for lumbar spinal stenosis: a self-administered, self-reported history questionnaire. *BMC Musculoskelet Disord*, 2007. 8: p. 102.
59. Lijmer, J.G., et al., Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*, 1999. 282(11): p. 1061-6.
60. Liao, K.P., et al., Anti-cyclic citrullinated peptide revised criteria for the classification of rheumatoid arthritis. *Ann Rheum Dis*, 2008. 67(11): p. 1557-61.
61. Taylor, W., et al., Classification criteria for psoriatic arthritis: development of new criteria from a large international study. *Arthritis Rheum*, 2006. 54(8): p. 2665-73.
62. van der Windt, D.A., et al., Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain. *Cochrane Database Syst Rev*, 2010(2): p. CD007431.

6

The association between lumbar disc degeneration and low back pain

The influence of age, gender, and individual radiographic features

Evelien I.T. de Schepper, Jurgen Damen, Joyce B.J. van Meurs, Abida Z. Ginai, Maria Popham, Albert Hofman, Bart W. Koes, Sita M. Bierma-Zeinstra

Abstract

Study design

Cross-sectional open population based study (nested in a prospective cohort study).

Objective

To explore the association of the different individual radiographic features, including osteophytes and disc space narrowing, with self-reported low back pain. Different definitions of lumbar disc degeneration with self-reported low back pain and disability were considered in a large open population sample. Furthermore, in order to disentangle the discrepancies in reported strength of the associations, we characterized the frequency of the different individual radiographic features of lumbar disc degeneration and definitions of lumbar disc degeneration, as well as their association with low back pain status, by age, gender and vertebral level.

Summary of background data

Currently within the literature, there have been no studies that have explored different definitions of lumbar disc degeneration and their association with low back pain within one study sample.

Methods

The intervertebral disc spaces (L1/2 to L5/S1) were evaluated for the presence and severity of anterior osteophytes and disc space narrowing using a semi-quantitative score (grade 0-3). Logistic regression was used to determine the association between these individual radiographic features of lumbar disc degeneration and different definitions of lumbar disc degeneration for low back pain.

Results

Lumbar radiographs were scored for 1204 men, and 1615 women. Osteophytes were the most frequent radiographic feature observed, with men having the greatest frequency. Disc space narrowing was more frequent in women than men. Both radiographic features increased in frequency with age.
Disc space narrowing appeared more strongly associated with low back pain than osteophytes, especially in men (odds ratio (OR) = 1.9; 95% confidence interval (CI): 1.4 to 2.8). Disc space narrowing at two or more levels appeared more strongly associated with low back pain than disc space narrowing at only one level (OR = 2.4; 95% CI: 1.6 to 3.4). After excluding level L5/S1, the strength of almost all associations increased.

Conclusions

We are the first to report different possible lumbar disc degeneration definitions and their associations with low back pain. Disc space narrowing at two or more levels appeared more strongly associated with low back pain than other radiographic features, especially after excluding level L5/S1.

Introduction

Back pain is one of the most common musculoskeletal complaints of the elderly, with a point prevalence of 26.9% in the Netherlands [1]. Van Tulder et al., [2] preformed a systematic review and reported that lumbar disc degeneration (LDD) could be a possible risk factor for back pain in adults, with odds ratios varying from 1.3 to 3.2. However, the review reported that the methodological quality of most of these studies was low. They also stated that the studies were difficult to compare due to difference in gender frequencies, age groups, settings, radiographic grading systems and definitions for LDD.

LDD is characterized radiologically by the presence of osteophytes, endplate sclerosis, and disc space narrowing. In 1993, Lane et al., presented a reliable grading system for these individual radiographic features (IRF) [3]. In a recent review [4], this grading system was recommended for use in epidemiologic studies, as their Interclass Correlation Coefficient (ICC) for inter-observer reliability were >0.60, with the exception of endplate sclerosis. There have been a number of recent studies that have used the classification of the IRF of disc degeneration, as defined in Lane et al., [5-7]. One of these studies described the occurrence of these separate features and their relationship with back pain in the open population, but only in a limited sample [6].

However, it is still unknown how to combine the IRF and how to define a clinically relevant definition for LDD. For example, currently there is no consensus about whether the lumbosacral disc should be scored. Some studies have included the lumbosacral level in their definition of LDD [8-10], while others have not [6, 11]. Currently within the literature, there have been no studies that have explored different definitions of LDD and their association with low back pain (LBP) within one study sample.

The purpose of this study was to explore the association of the different IRF, including osteophytes and disc space narrowing, with self-reported LBP. Different definitions of LDD with self-reported LBP and disability were considered in a large open population sample. Furthermore, in order to disentangle the discrepancies in reported strength of the associations, we characterized the frequency of the different IRF of LDD and definitions of LDD, as well as their association with LBP status, by age, gender and vertebral level.

Materials and methods

Study population

The data for this study originate from data of the Rotterdam Study, an open population prospective cohort of people aged 55 years and older. The study design has been described previously [12]. The baseline measurements were conducted between 1990 and 1993. The focus was on neurogeriatric, cardiovascular, ophthalmologic and locomotor diseases. At baseline, trained interviewers performed an extensive home interview on demographic characteristics, medical history, risk factors for chronic diseases and medication use. Radiographs were taken at the research centre at baseline. In total, 7983 participants were examined, however, for feasibility reasons, only lumbar radiographs of 2819 participants were scored. These participants were selected on availability of radiograph data for a follow-up measurement 6.6 years later.

Radiographic scoring

Lumbar lateral radiographs were scored by a single observer trained by a radiologist for the presence of the individual radiographic features of disc degeneration. The observer was blinded to clinical characteristics of the participants. Each vertebral level from L1/2 to L5/S1 was reviewed for the presence and severity of osteophytes (anterior) and vertebral narrowing, using the Lane et al. atlas [3, 4] . In this atlas grade 0 = none; grade 1 = mild; grade 2 = moderate; and grade 3 = severe. The lumbosacral disc space was defined as narrowed when its height was less than that of the disc space between the third and fourth lumbar vertebrae. This is due to a normal progression of increasing disc-space height from the third and fourth to the fourth and fifth lumbar vertebrae, and then a relative narrowing of the height of the lumbosacral disc space. Sclerosis was not scored because of the earlier reported low ICC for this feature [3].
Inter-observer reproducibility was assessed by a second independent observer who evaluated a random selection of 140 (5%) X-rays. The ICC was 0.83 for osteophytes and 0.77 for vertebral narrowing, indicating good reproducibility.

Back pain and disability

Back pain was determined from interviewing the participants during the home visits. Participants were asked “Did you have complaints of the low back during the last month?”. LBP was defined to be present if the answer

was positive. Participants were also asked “What is the duration of the present low back complaints?”. We defined chronic LBP to be present if the duration of the LBP was more than one year. In this way the definition chronic LBP included long lasting chronic complaints.
Participants also visited the research center, where radiographs were carried out. Height and weight were measured with participants wearing indoor clothing and without shoes. Body mass index (BMI) was calculated as weight in kilograms divided by length in meters squared (kg/m^2). Dual energy radiograph absorptiometry was used to assess the bone mineral density (BMD) at the femoral neck. Stanford Health Assessment Questionnaire (HAQ) was used to assess disability. A mean score ≥ 0.5 was used to indicate a moderate to severe disability [13].

Statistical Analyses
We defined disc space narrowing to be present if the grade was mild, moderate, or severe (grade ≥1). Because of the small proportion of subjects without osteophytes, we used a higher cutoff value for this feature. We defined osteophytes to be present if the grade was moderate or severe (grade ≥2). Using these definitions we calculated the prevalence of the IRF by vertebral level (L1/2 to L5/S1), age and gender.
To analyze data we defined the lower back as one unit with five sub-joints. Subsequently, we calculated the frequency of each of the IRF at the lower back with the question: “Is there somewhere in the lower back a certain grade of narrowing/osteophytes?”.
We characterize three possible LDD definitions, “narrowing”, “osteophytes”, and “both”. “Narrowing” is positive when there is a grade ≥ 1 narrowing at two or more levels and “osteophytes” is positive when there is a grade ≥ 2 osteophytes at two or more levels. When “narrowing” and “osteophytes” both are positive, it is assigned “both”. To explore the role of the lumbosacral disc space, we also investigate the three LDD definitions after excluding the vertebral level L5/S1. These were classified “narrowing 1-4”, “osteophytes 1-4” and “both 1-4”.
To explore the association between the IRF, the LDD definitions and LBP, LBP was used as the dependent variable with adjustments made for age and gender. The assessments of the association were adjusted for BMI, as this variable has been reported to be associated with both LBP and some of the individual LDD features [5, 9, 14, 15]. The associations were also adjusted for BMD, even though this variable has not been reported to be associated with LBP [16]. However, in our data BMD was shown to be

associated with LBP in men but not in women. The associations were not adjusted for smoking and education, as these variables were shown to be not associated with LBP. The results of these analyses are expressed as odds ratios (OR) with 95% confidence intervals (CI), stratified for gender and age groups. The same methods were used to explore the association between the IRF and chronic LBP. The associations with disability were additionally corrected for hip OA and knee OA (defined by Kellgren and Lawrence grade ≥ 2); Kellgren and Lawrence data of the knee OA were only available for 50% of the participants. Therefore, these latter analyses were limited to half of the sample. Finally, the independency of the different features in the association with LBP was checked; besides the correcting variables, both features were simultaneously included in the model. The model included the two most favorable definitions for osteophytes and narrowing ("narrowing 1-4" and "osteophytes 1-4"), and subsequently two summary scores as the grade of the number of levels affected by osteophytes or narrowing.
Statistical analysis was performed using SPSS version 15 (SPSS Inc, Chicago, USA).

Results

Subject characteristics
Baseline characteristics are shown in Table 1. There were 1204 men (mean age 65.3 years, standard deviation (SD) 6.4) and 1615 women (mean age 65.9 years, SD 6.8). LBP during the last month was reported more often by women than men (326 (20.2%) vs. 173 (14.4%) $p < 0.05$) (Table 1). Chronic LBP was reported in 84% of the current LBP cases and was also more often reported by women (280 (17.3%) vs. 140 (11.6%) $p < 0.05$).

Influence of gender and vertebral level
The prevalence of the IRF in men and women is shown in Table 1. Osteophytes were the most frequent observed radiographic feature and were more common in men than women (95% vs. 91%; $p<0.05$). Disc space narrowing was more frequent in women than men (65% vs. 53%; $p<0.05$). In terms of their distribution by vertebral level, narrowing was more frequent at the lower lumbar disc levels.

Table 1 Frequency of osteophytes and disc space narrowing in men and women.

	Men, N = 1204	**Women, N = 1615**	**All, N = 2819**	**LBP, N = 499**
Age (years) Mean ± SD	65.3 ± 6.4	65.9 ± 6.8	65.7 ± 6.6	65.4 ± 6.7
Body mass index (BMI) Mean ± SD	25.9 ± 2.9	26.6 ± 3.8	26.3 ± 3.5	26.4 ± 3.6
Bone mineral density femoral neck (BMD) Mean ± SD	0.89 ± 0.12	0.82 ± 0.13	0.85 ± 0.13	0.85 ± 0.13
Low back pain (%)†	173 (14.4)	326 (20.2)	499 (17.7)	499 (100)
Chronic low back pain (%)‡	140 (11.6)	280 (17.3)	420 (14.9)	420 (84.2)
Osteophytes low back (%)				
Grade ≥ 1	1148 (95.3)	1467 (90.8)	2615 (92.8)	469 (94.0)
Grade ≥ 2	832 (69.1)	929 (57.5)	1761 (62.5)	323 (64.7)
Grade 3	536 (44.5)	505 (31.3)	1041 (36.9)	199 (39.9)
Narrowing low back (%)				
Grade ≥ 1	637 (52.9)	1048 (64.9)	1685 (59.8)	335 (67.1)
Grade ≥ 2	286 (23.8)	525 (32.5)	811 (28.8)	173 (34.7)
Grade 3	40 (3.3)	107 (6.6)	147 (5.2)	37 (7.4)
Osteophytes ≥2 (%)				
L1-2	282 (23.4)	297 (18.4)	579 (20.5)	112 (22.4)
L2-3	347 (28.8)	404 (25.0)	751 (26.6)	155 (31.1)
L3-4	428 (35.5)	364 (22.6)	792 (28.1)	155 (31.1)
L4-5	403 (33.5)	354 (21.9)	757 (26.9)	145 (29.1)
L5-S1	312 (25.9)	303 (18.8)	615 (21.8)	121 (24.2)
Narrowing ≥1 (%)				
L1-2	107 (8.9)	201 (12.5)	308 (10.9)	71 (14.2)
L2-3	135 (11.3)	307 (19.0)	442 (15.7)	115 (23.0)
L3-4	153 (12.7)	342 (21.1)	495 (17.6)	136 (27.3)
L4-5	268 (22.2)	526 (32.6)	794 (28.2)	187 (37.5)
L5-S1	408 (34.0)	662 (41.0)	1070 (38.0)	214 (42.9)

† Low back pain: complaints of the low back during last month
‡ Chronic low back pain: duration present low back complaints > 1 year

Association with back pain

Table 2 shows the association between LDD and low back pain, adjusted for age, gender, BMI and BMD. The presence of disc space narrowing grade ≥ 1 and ≥ 2 was significantly associated with LBP in the last month, only in men (Nar ≥ 1 OR = 1.9; 95% CI: 1.4 to 2.8 and Nar ≥ 2 OR = 1.6; 95% CI: 1.1 to 2.4) (Table 2). The presence of osteophytes grade ≥ 2 was not significantly associated with back pain in men or women.

"Narrowing" was associated with LBP in men and women (men OR = 2.4; 95% CI: 1.6 to 3.4 and women OR = 1.7; 95% CI: 1.3 to 2.3). "Osteophytes"

Table 2. Association between disc degeneration and low back pain.

	Men, N = 1204			Women, N = 1615			All, N = 2819		
	N (%)	LBP OR (95% CI)	Chronic LBP OR (95% CI)	N (%)	LBP OR (95% CI)	Chronic LBP OR (95% CI)	N (%)	LBP OR (95% CI)	Chronic LBP OR (95% CI)
Ost ≥2	832 (69.1)	1.3 (0.9 - 1.9)	1.4 (0.9 - 2.1)	929 (57.5)	1.1 (0.9 - 1.5)	1.2 (0.9 - 1.5)	1761 (62.5)	1.2 (1.0 - 1.5)	1.2 (1.0 - 1.5)
Nar ≥1	637 (52.9)	1.9 (1.4 - 2.8)**	2.2 (1.5 - 3.3)**	1048 (64.9)	1.1 (0.9 - 1.5)	1.3 (1.0 - 1.7)	1685 (59.8)	1.4 (1.1 - 1.7)**	1.6 (1.2 - 2.0)**
Nar ≥2	286 (23.8)	1.6 (1.1 - 2.4)*	1.9 (1.3 - 2.8)**	525 (32.5)	1.2 (0.9 - 1.6)	1.3 (1.0 - 1.7)	811(28.8)	1.3 (1.1 - 1.7)**	1.5 (1.2 - 1.8)**
“Narrowing”	277 (23.0)	2.4 (1.6 - 3.4)**	3.0 (2.0 - 4.4)**	553 (34.2)	1.7 (1.3 - 2.3)**	1.9 (1.4 - 2.5)**	830 (29.4)	2.0 (1.6 - 2.4)**	2.2 (1.8 - 2.8)**
“Osteophytes”	518 (43.0)	1.5 (1.1 - 2.2)*	1.9 (1.3 - 2.7)**	481 (29.8)	1.5 (1.2 - 2.0)**	1.5 (1.1 - 2.0)**	999 (35.4)	1.5 (1.2 - 1.9)**	1.6 (1.3 - 2.0)**
“Both”	203 (16.9)	2.2 (1.4 - 3.2)**	2.8 (1.8 - 4.2)**	300 (18.6)	1.9 (1.4 - 2.6)**	1.9 (1.4 - 2.7)**	503 (17.8)	2.0 (1.6 - 2.5)**	2.2 (1.7 - 2.8)**
Ost ≥2 1-4	746 (62.0)	1.4 (1.0 - 2.1)*	1.6 (1.1 - 2.4)*	809 (50.1)	1.2 (0.9 - 1.5)	1.2 (0.9 - 1.6)	1555 (55.2)	1.3 (1.0 - 1.6)*	1.3 (1.1 - 1.7)*
Nar ≥1 1-4	421 (35.0)	2.2 (1.6 - 3.1)**	2.6 (1.8 - 3.8)**	770 (47.7)	1.3 (1.0 - 1.7)*	1.4 (1.1 - 1.9)*	1191 (42.2)	1.6 (1.3 - 2.0)**	1.8 (1.4 - 2.2)**
Nar ≥2 1-4	137 (11.4)	2.5 (1.6 - 4.0)**	3.0 (1.9 - 4.8)**	290 (18.0)	1.5 (1.1 - 2.0)*	1.5 (1.1 - 2.1)*	427 (15.1)	1.8 (1.4 - 2.3)**	1.9 (1.5 - 2.5)**
“Narrowing 1-4”	162 (13.5)	2.4 (1.5 - 3.6)**	3.0 (1.9 - 4.8)**	363 (22.5)	2.0 (1.5 - 2.7)**	2.2 (1.7 - 3.0)**	525 (18.6)	2.1 (1.7 - 2.7)**	2.5 (1.9 - 3.2)**
“Osteophytes 1-4”	429 (35.6)	1.2 (0.8 - 1.7)	1.4 (1.0 - 2.1)	410 (25.4)	1.5 (1.1 - 2.0)**	1.4 (1.0 - 1.9)*	839 (29.8)	1.4 (1.1 - 1.7)**	1.4 (1.1 - 1.8)**
“Both 1-4”	118 (9.8)	1.9 (1.2 - 3.1)*	2.5 (1.5 - 4.1)**	210 (13.0)	1.7 (1.2 - 2.4)**	1.6 (1.1 - 2.4)**	328 (11.6)	1.8 (1.3 - 2.3)**	1.9 (1.4 - 2.5)**

Adjusted for age, gender, BMI and BMD

* p = < 0.05 **; p = < 0.01

OR odds ratio CI; confidence interval LBP; low back pain

Ost ≥2	osteophytes grade ≥ 2 in the low back	Ost ≥2 1-4	Ost ≥2 from L1/L2 to L4/L5
Nar ≥1	narrowing grade ≥ 1 in the low back	Nar ≥1 1-4	Nar ≥1 from L1/L2 to L4/L5
Nar ≥2	narrowing grade ≥ 2 in the low back	Nar ≥2 1-4	Nar ≥2 from L1/L2 to L4/L5
“Narrowing”	grade ≥ 1 narrowing at two or more levels	“Narrowing 1-4”	“Narrowing” from L1/L2 to L4/L5
“Osteophytes”	grade ≥ 2 osteophytes at two or more levels	“Osteophytes 1-4”	“Osteophytes” from L1/L2 to L4/L5
“Both”	“Narrowing” and “Osteophytes” positive	“Both 1-4”	“Both” from L1/L2 to L4/L5

was equally strong associated with back pain in men and women (men OR = 1.5; 95% CI: 1.1 to 2.2 and women OR = 1.5; 95% CI: 1.2 to 2.0). "Both" also showed an association with back pain for both men and women (men OR = 2.2; 95% CI: 1.4 to 3.2 and women OR 1.9; 95% CI: 1.4 to 2.6). The strength of almost all the associations increased for chronic low back pain, especially for men ("narrowing" OR = 3.0; 95% CI: 2.0 to 4.4).
After including both features in the model, the presence of osteophytes was not significantly associated with back pain in men or women ("osteophytes 1-4" men OR = 1.0; 95% CI: 0.7 to 1.4 and women OR = 1.2; 95% CI: 0.9 to 1.6).

Lumbosacral level
The strength of the associations including disc space narrowing increased by excluding level L5/S1, particularly for "narrowing1-4" in both men and women (men OR = 2.4; 95% CI: 1.5 to 3.6 and women OR = 2.0; 95% CI: 1.5 to 2.7).
The presence of disc space narrowing, with grade ≥ 1 and ≥ 2 was associated with LBP in both men and women, after excluding level L5/S1 (Table 2). The strength of the association also increased with increasing severity of disc space narrowing (Nar ≥ 1 1-4; OR 2.2; 95% CI: 1.6 to 3.1 and Nar ≥ 2 1-4; OR 2.5: 95% CI: 1.6 to 4.0).

Influence of age
The prevalence of osteophytes and disc space narrowing increased with age for both men and women (Table 3). In men, "narrowing" was more greatly associated with LBP in the age groups 55-59 and 60-64 (55-59 OR = 3.5; 95% CI: 1.7 to 7.5 and 60-64 OR = 4.1; 95% CI: 2.2 to 7.8), while in women, "narrowing" was highly associated with LBP in the age groups 65-69 and 70-74 (65-69 OR = 2.0; 95% CI: 1.1 to 3.5 and 70-74 OR 3.1; 95% CI: 1.7 to 5.8).

Disability
After adjusting for age, sex, BMI, BMD, and radiologic osteoarthritis of the hip and knee, "narrowing" was associated with disability (men OR = 2.1; 95% CI: 1.1 to 3.7; women OR = 2.0; 95% CI: 1.4 to 2.9; all OR = 1.9; 95% CI: 1.4 to 2.6). After excluding level L5/S1, the strength of the association remained stable (all OR = 2.0; 95% CI: 1.4 to 2.8). After including LBP in the model, the strength of the association diminished only slightly (all OR = 1.7; 95% CI: 1.2 to 2.4).

Table 3. Association between individual radiographic features of lumbar disc degeneration, low back pain and age groups in men and women.

Age, yrs	N (%)	LBP (%)	Ost ≥2 (%)	LBP OR (95%CI)	Nar ≥1 (%)	LBP OR (95%CI)	"Narrowing" (%)	LBP OR (95% CI)	"Narrowing" (%)	LBP OR (95% CI)
Men, N = 1204										
55 - 59	271 (22.5)	43 (15.9)	163 (60.1)	1.5 (0.7 - 3.1)	118 (43.5)	2.9 (1.4 - 5.8)**	45 (16.6)	3.5 (1.7 - 7.5)**	25 (9.2)	6.4 (2.7 - 15.6)**
60 - 64	352 (29.2)	59 (16.8)	241 (68.1)	2.3 (1.1 - 4.7)*	184 (52.3)	2.0 (1.1 - 3.6)*	64 (18.2)	4.1 (2.2 - 7.8)**	28 (8.0)	2.7 (1.1 - 6.3)*
65 - 69	301 (25.0)	44 (14.6)	211 (70.1)	0.8 (0.4 - 1.8)	160 (53.2)	2.2 (1.1 - 4.4)*	69 (22.9)	2.0 (1.0 - 4.1)	43 (14.3)	2.2 (1.0 - 5.1)
70 - 74	179 (14.9)	16 (8.9)	133 (74.3)	0.8 (0.2 - 2.7)	105 (58.7)	0.6 (0.2 - 1.7)	56 (31.3)	1.1 (0.4 - 3.4)	37 (20.7)	1.4 (0.4 - 4.7)
75 +	101 (8.4)	11 (10.9)	84 (83.2)	0.3 (0.1 - 1.2)	70 (69.3)	1.3 (0.3 - 5.9)	43 (42.6)	0.4 (0.1 - 1.8)	29 (28.7)	0.3 (0.0 - 2.5)
Women, N = 1615										
55 - 59	364 (22.5)	79 (21.7)	189 (51.9)	1.2 (0.7 - 1.9)	197 (54.1)	0.9 (0.5 - 1.5)	70 (19.2)	1.8 (1.0 - 3.3)	45 (12.4)	2.0 (1.0 - 4.0)
60 - 64	437 (27.1)	79 (18.1)	237 (54.2)	1.4 (0.9 - 2.4)	261 (59.7)	1.7 (1.0 - 3.0)	123 (28.1)	1.5 (0.9 - 2.7)	67 (15.3)	1.8 (1.0 - 3.5)
65 - 69	357 (22.1)	64 (17.9)	210 (58.8)	1.0 (0.6 - 1.7)	249 (69.7)	0.8 (0.5 - 1.5)	141 (39.5)	2.0 (1.1 - 3.5)*	88 (24.6)	2.0 (1.1 - 3.6)*
70 - 74	280 (17.3)	68 (24.3)	174 (62.1)	1.0 (0.6 - 1.9)	201 (71.8)	2.5 (1.2 - 5.2)*	120 (42.9)	3.1 (1.7 - 5.8)**	83 (29.6)	3.1 (1.7 - 5.8)**
75 +	177 (11.0)	36 (20.3)	119 (67.2)	1.2 (0.5 - 2.8)	140 (79.1)	0.6 (0.2 - 1.5)	99 (55.9)	0.8 (0.4 - 1.8)	80 (45.2)	1.4 (0.6 - 3.1)
All, N = 2819										
55 - 59	635 (22.5)	122 (19.2)	352 (55.4)	1.3 (0.8 - 1.9)	315 (49.6)	1.4 (0.9 - 2.1)	115 (18.1)	2.3 (1.4 - 3.7)**	70 (11.0)	3.0 (1.7 - 5.2)**
60 - 64	789 (28.0)	138 (17.5)	478 (60.6)	1.7 (1.1 - 2.6)*	445 (56.4)	1.9 (1.2 - 2.8)**	187 (23.7)	2.3 (1.5 - 3.5)**	95 (12.0)	2.1 (1.2 - 3.5)**
65 - 69	658 (23.3)	108 (16.4)	421 (64.0)	0.9 (0.6 - 1.4)	409 (62.2)	1.2 (0.8 - 2.0)	210 (31.9)	2.0 (1.3 - 3.1)**	131 (19.9)	2.1 (1.3 - 3.4)**
70 - 74	459 (16.3)	84 (18.3)	307 (66.9)	1.0 (0.6 - 1.7)	306 (66.7)	1.6 (0.9 - 2.9)	176 (38.3)	2.5 (1.5 - 4.2)**	120 (26.1)	2.6 (1.5 - 4.5)**
75 +	278 (9.9)	47 (16.9)	203 (73.0)	0.8 (0.4 - 1.7)	210 (75.5)	0.7 (0.3 - 1.6)	142 (51.1)	0.7 (0.4 - 1.4)	109 (39.2)	1.4 (0.7 - 3.0)

Adjusted for gender, BMI and BMD

*	p = < 0.05	OR	odds ratio	Ost ≥2	osteophytes ≥ 2 in the low back
**	p = < 0.01	CI	confidence interval	Nar ≥1	narrowing ≥ 1 in the low back
LBP	low back pain	"Narrowing"	grade ≥1 narrowing at two or more levels	Narrowing 1-4"	"Narrowing" from L1/L2 to L4/L5

Discussion

We are the first to report in one paper, multiple LDD definitions and their associations with low back pain, for the separate genders and discreet age groups. In this study, disc space narrowing appeared to be more strongly associated with LBP than osteophytes, especially in men. Disc space narrowing at two or more levels appeared more strongly associated with LBP than disc space narrowing at only one level. The strength of the associations increased with chronic LBP. The majority of the associations were strengthened by excluding level L5/S1.
The most frequently observed radiographic feature of LDD was osteophytes, with greater frequency in men than in women. Narrowing, however, was more common in women than in men and was also shown to be more frequent at the lower disc levels. Both IRF increased in frequency with age.

How do these findings compare with those of other studies?
Data from many studies suggest an association with LDD and low back pain with odds ratio varying from 1.3 to 3.2 [2]. However, we are not aware of data from population-based samples that have investigated the association of different definitions of lumbar disc degeneration with self-reported low back pain. MacGregor et al.,[17] preformed a study, using MRI scans to assess risk factors associated with severe back pain. They investigated a number of features including; disc height, signal change, disc bulge and anterior osteophytes, and made a sum score for all features together. This sum was associated with severe back pain. However they did not state which features had the highest predictive capability. Some studies suggest an association between osteophytes and LBP [6, 8, 10] and some studies suggest an association between disc space narrowing and LBP [6, 8, 9, 18]. Our data confirms the association between LBP and disc space narrowing. In addition, our data suggest an association with osteophytes, only when a more specific definition ("osteophytes") is used. However, when osteophytes and disc space narrowing were both included in the model, there was no association with osteophytes anymore. Therefore osteophytes do not have an independent association with low back pain and seem therefore an inferior derivate from disc space narrowing.
Some studies suggest that the strength of the association between LBP and disc space narrowing grows with increasing severity of disc space narrowing [6, 9]. Our data confirm this, but only when L5/S1 is excluded.

Further, our data indicate that the association between LBP and disc space narrowing increases when a greater number of levels are affected.

The explanation for the stronger association between LBP and disc space narrowing compared with the presence of osteophytes is unknown. It is possible that the reduction of space between the vertebrae as a consequence of the degenerative disc is more likely to lead to increased pressure on facet joints and spinal ligaments.
The explanation for the stronger association between back pain and disc space in men compared with the association in women is also unknown. It is possible that even though women reported LBP more often, only a small proportion of the complaints are due to LDD, whereas other factors determine the feeling of pain. Men and women could also report pain differently therefore effecting the association between back pain, disc space narrowing and gender. Cecchi et al., showed that women presented with significantly more severe pain than men [19].
A possible explanation for the stronger association between LBP and disc space narrowing, excluding level L5/S1, is the possible overrating of the narrowing grade of the lumbosacral disc. The height of the lumbosacral disc is difficult to score due to its narrowed height relative to disc L4/L5. The lumbosacral disc is also different in appearances among different individuals, independently of disease [8, 20]. Therefore, by using the lumbar disc definition "narrowing 1-4", the inconsistency of the grading scores at this level is ruled out. Furthermore, some differences in the reported associations in the prior studies can be explained by our stratified results. Possible explanations for relatively low odds ratios previously reported could be due to the use of a young age group [8], the use of women only [9] and scoring of the lumbosacral level [8-10].

Our data confirms the findings from recent population based radiographic surveys showing a greater frequency and severity of osteophytes in men than in women [6, 10]. A possible explanation for the greater frequency in men is the higher BMD in men. However, after including BMD in the model, although less explicit, men still show a greater frequency and severity of osteophytes.
Surprisingly our data suggest a greater frequency and severity of narrowing in women than in men. We found no explanation for this finding so far. Our data confirms that the prevalence of osteophytes and disc space narrowing increases with age in both men and women [6].

Our study had several advantages. It was population based with a relatively high number of subjects. We used a semi-quantitative score, using standard radiographs, to characterize the presence and severity of LDD. Assessment of the radiographs was taken without knowledge of the questionnaire data, and so errors in classification are likely to have been non-directional. However, there are several limitations in our study that need to be considered when interpreting the results. There could be selection bias in favor of relatively healthy participants. The participants in the present study had to be mobile enough to visit the research center for radiograph examination, both for the baseline and follow up appointments (mean 6.6 years) [21]. In other words, patients with the most severe symptoms were most likely not included, but this may be inevitable in long term prospective cohort studies.

The results of the present study may be flawed by the decision not to use a frontal lumbar radiographs. Therefore, there is a possibility that a single sided disc space narrowing or a lateral osteophyte is missed. Also, on a lateral radiograph the facet joints cannot be judged. However, clinical insights indicate that choosing between both of them, a lateral lumbar radiographs gives more information about disc space narrowing and osteophytes.

In literature, others have chosen three months or even six months as the dividing line between acute and chronic pain [22]. However, we defined chronic LBP to be present when the duration of the LBP was more than one year. In this way the definition chronic LBP included long lasting chronic complaints with long lasting impact on one's life. When we defined chronic LBP to be present when the duration of the LBP was more than six months, the OR of the associations diminished with 0.1 to 0.2.

What are the implications of these findings for researchers and clinicians?

From our data, a useful case definition for LDD can be deduced; specifically disc space narrowing at two or more levels from L1/2 to L4/L5. This definition shows the strongest relationship with LBP and represents a more generalized form of LDD. As a result it might be a promising clinically relevant phenotype in genetic and epidemiologic LDD research.

Our data provides evidence for a moderate association between disc space narrowing and LBP. This association is only slightly less than the association of pain and radiological knee osteoarthritis [23] and even slightly more than the association of pain and radiological hand osteoarthritis [24] in the same population sample.

The most important aspect of our data is that disc space narrowing at two or more levels is even more related to chronic LBP. The ability of LDD in predicting LBP at the follow up period was unfortunately not possible to investigate, as no questions about LBP specifically were asked at the follow up visit.

Conclusion

In conclusion, our data provides evidence for an association between disc space narrowing and LBP especially in men, with the association increasing, with increasing numbers of affected intervertebral disc spaces. Furthermore, our data highlights the frequent occurrence of IRF, as well as the increased frequency in age, of the IRF of LDD in population samples of men and women.

References

1. Picavet HSJ, v.G.H., Schouten JSAG, Klachten van het bewegingsapparaat in de Nederlandse bevolking; prevalenties, consequenties en risicogroepen. *RIVM*, 2000(rapport 266807002).
2. van Tulder, M.W., et al., Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine*, 1997. 22(4): p. 427-34.
3. Lane, N.E., et al., Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol*, 1993. 20(11): p. 1911-8.
4. Kettler, A. and H.J. Wilke, Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. *Eur Spine J*, 2006. 15(6): p. 705-18.
5. Yoshimura, N., et al., Epidemiology of chronic disc degeneration and osteoarthritis of the lumbar spine in Britain and Japan: a comparative study. *J Rheumatol*, 2000. 27(2): p. 429-33.
6. Pye, S.R., et al., Radiographic features of lumbar disc degeneration and self-reported back pain. *J Rheumatol*, 2004. 31(4): p. 753-8.
7. Hassett, G., et al., Risk factors for progression of lumbar spine disc degeneration: the Chingford Study. *Arthritis Rheum*, 2003. 48(11): p. 3112-7.
8. Frymoyer, J.W., et al., Spine radiographs in patients with low-back pain. An epidemiological study in men. *J Bone Joint Surg Am*, 1984. 66(7): p. 1048-55.
9. Symmons, D.P., et al., A longitudinal study of back pain and radiological changes in the lumbar spines of middle aged women. II. Radiographic findings. *Ann Rheum Dis*, 1991. 50(3): p. 162-6.
10. Biering-Sorensen, F., et al., The relation of spinal x-ray to low-back pain and physical activity among 60-year-old men and women. *Spine*, 1985. 10(5): p. 445-51.
11. Oishi, Y., et al., Lack of association between lumbar disc degeneration and osteophyte formation in elderly Japanese women with back pain. *Bone*, 2003. 32(4): p. 405-11.
12. Hofman, A., et al., The Rotterdam Study: objectives and design update. *Eur J Epidemiol*, 2007. 22(11): p. 819-29.
13. Bruce, B. and J.F. Fries, The Stanford Health Assessment Questionnaire: dimensions and practical applications. *Health Qual Life Outcomes*, 2003. 1: p. 20.
14. van Saase, J.L., et al., Osteoarthritis and obesity in the general population. A relationship calling for an explanation. *J Rheumatol*, 1988. 15(7): p. 1152-8.

15. Pye, S.R., et al., Influence of weight, body mass index and lifestyle factors on radiographic features of lumbar disc degeneration. *Ann Rheum Dis*, 2007. 66(3): p. 426-7.
16. Hangai, M., et al., Factors associated with lumbar intervertebral disc degeneration in the elderly. *Spine J*, 2008. 8(5): p. 732-40.
17. MacGregor, A.J., et al., Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins. *Arthritis Rheum*, 2004. 51(2): p. 160-7.
18. Torgerson, W.R. and W.E. Dotter, Comparative roentgenographic study of the asymptomatic and symptomatic lumbar spine. *J Bone Joint Surg Am*, 1976. 58(6): p. 850-3.
19. Cecchi, F., et al., Epidemiology of back pain in a representative cohort of Italian persons 65 years of age and older: the InCHIANTI study. *Spine*, 2006. 31(10): p. 1149-55.
20. Henry, G.W., I.J. Larsen, and S.F. Stewart, The roentgenologic criteria for appraising the human back as an economic asset or liability. *Am J Roentgenol Radium Ther Nucl Med*, 1958. 79(4): p. 658-72.
21. Reijman, M., et al., Validity and reliability of three definitions of hip osteoarthritis: cross sectional and longitudinal approach. *Ann Rheum Dis*, 2004. 63(11): p. 1427-33.
22. Bonica, J., The management of pain. 1990, Philadelphia: Lea and Febiger.
23. Claessens, A.A., et al., Do clinical findings associate with radiographic osteoarthritis of the knee? *Ann Rheum Dis*, 1990. 49(10): p. 771-4.
24. Dahaghin, S., et al., Prevalence and pattern of radiographic hand osteoarthritis and association with pain and disability (the Rotterdam study). *Ann Rheum Dis*, 2005. 64(5): p. 682-7.

7

Association between spinal morning stiffness and lumbar disc degeneration

The Rotterdam Study

J. Scheele, E.I.T. de Schepper,
J.B.J. van Meurs, A. Hofman, B.W. Koes,
P.A.J. Luijsterburg,
S.M.A. Bierma-Zeinstra

Abstract

Objective

To explore the associations between spinal morning stiffness and lumbar disc degeneration (LDD).

Design

Data from a cross-sectional general-population-based study (Rotterdam Study-I) were used. Intervertebral disc spaces and osteophytes of people aged ≥55 years were scored on lumbar lateral radiographs (L1-2 through L5-S1 was scored). Logistic regression analysis was used to explore associations between spinal morning stiffness and two definitions of LDD (i.e. 'narrowing' and 'osteophytes'). Spinal morning stiffness combined with low back pain and its association with LDD was also analyzed. Similar analyses were performed for knee and hip pain, morning stiffness in the legs, and radiographic knee and hip osteoarthritis (OA) in order to compare these associations with those of LDD. All analyses were adjusted for age, gender, and body mass index (BMI).

Results

Lumbar lateral radiographs were scored for 2819 participants. Both definitions of LDD were associated with spinal morning stiffness: adjusted odds ratio (aOR) 1.3; 95% confidence interval (CI): 1.1-1.6 for 'osteophytes' and aOR 1.8; 95% CI: 1.4-2.2 for 'narrowing'. Both the odds ratios increased when spinal morning stiffness was combined with low back pain: aOR 1.5; 95% CI: 1.1-2.0 for 'osteophytes' and aOR 2.5; 95% CI: 1.9-3.4 for 'narrowing'. When morning stiffness in the legs was combined with knee or hip pain, the associations with radiographic knee or hip OA were: aOR 3.0; 95% CI: 2.1-4.1 for knee OA and aOR 3.1; 95% CI: 1.9-5.0 for hip OA.

Conclusions

Reported spinal morning stiffness is associated with LDD. The associations increased when we combined spinal morning stiffness with low back pain. The magnitude of the association for the definition 'narrowing' is similar to the association between morning stiffness in the legs and knee or hip OA.

Introduction

Low back pain is a major health problem, also in the elderly. It is the most reported pain site of all musculoskeletal complaints [1]. Low back pain is often defined as pain possibly with muscle tension or stiffness, localized below the costal margin and above the inferior gluteal folds, with or without radiating leg pain [2]. Since patients with non-specific low back pain are not only a large but also a very heterogeneous group regarding etiology, prognosis and susceptibility to treatment, it is important to identify subgroups within this population. Low back pain patients with symptoms due to lumbar disc degeneration (LDD) or lumbar osteoarthritis (OA) could be such a subgroup, and clinical symptoms associated with radiographic LDD may help identify those with symptoms due to LDD or lumbar OA in clinical practice.

An association between radiographic LDD and low back pain has been reported in several studies [3-7]. The study of de Schepper et al. compared associations between different definitions of LDD and low back pain [3]. They found an association for the definition based on the presence of disc space narrowing, as well as for the definition based on the presence of osteophytes [3].

Although there are no official classification criteria for LDD, it is often characterized by narrowing of the disc space and the presence of osteophytes, seen at the lumbar radiograph [4]. Disk degeneration is associated with and often precedes facet joint OA [8-10]. Although LDD cannot be defined as real OA because the facet joints are the only synovial joints in the spine, LDD is often used as a proxy for OA of the spine, in particular when imaging (preferably with magnetic resonance imaging) of the synovial joints is not available. OA of the knee and hip already has clinical classification criteria, described by the American College of Rheumatology (ACR). The ACR criteria describe that, besides pain, morning stiffness is an important criteria for hip and knee OA [11, 12].

Therefore, the present study explores the association between: 1) spinal morning stiffness and LDD, and 2) spinal morning stiffness in combination with low back pain and LDD, cross-sectional in a large general population study. These associations are also compared with the associations between morning stiffness in the legs, and knee or hip OA.

Methods

Study population

This study used data from the Rotterdam Study, a general population prospective cohort study of people aged 55 years and older living in Rotterdam (The Netherlands). All inhabitants of Ommoord, a district of the city Rotterdam, aged 55 years and older (n= 10,215) were invited to participate in this study. In total, 7983 adults participated in the baseline measurements (78% of the invited inhabitants) [13]. The detailed study design has been described elsewhere [13, 14]. The present study used the baseline measurements (RS I-1) which were collected in 1990-1993, and included a home interview and radiographs made in a research center in the participant's district. The Medical Ethics Committee of Erasmus Medical Center approved the protocol of the Rotterdam Study. The present study consisted of a random selection of 2819 participants with spinal radiographs available at both baseline and at 6.6 years follow-up, as described in a previous study [3].

Radiographs

The lumbar spine levels L1-2 through L5-S1 were scored on the lateral lumbar radiograph for the presence and severity of osteophytes (anterior) and disc space narrowing, using the system of Lane et al. [15]. This system grades both osteophytes and disc space narrowing on a scale from 0-3, in which 0 = none, 1 = mild, 2 = moderate, and 3 = severe. The Lane atlas contains lumbar radiograph in which the different grades of osteophytes and narrowing are illustrated. Disc space narrowing was scored if the height between the lumbar vertebrae was different from the normal progression of the spine. The Lane atlas is one of the systems recommended in a recent review on existing grading scales [16].

All spinal radiographs were scored by a single reader [EdS], who was trained to score the radiographs and blinded to the participants' clinical data. A random selection of spinal radiographs (140; 5%) was evaluated by another trained reader to obtain the interobserver reproducibility. The intraclass correlation coefficient (ICC) was 0.83 for scoring osteophytes and 0.77 for scoring disc space narrowing, which indicates a good reproducibility [3].

An earlier report of the Rotterdam Study [3] analyzed the association between different radiographic features of LDD and low back pain. They concluded that the association increased after excluding level L5-S1 from

the analysis, and when disc space narrowing or osteophytes were present at two or more vertebral levels [3]. Disc space narrowing of the lumbosacral disc is also more difficult to score due to its narrow height and because the variable height of a normal disc at this level makes it difficult to establish pathology [17, 18]. We used the two different definitions of LDD proposed in the study of de Schepper et al., i.e. 'narrowing' and 'osteophytes'. 'Narrowing' is defined as disc space narrowing (grade ≥1) at two or more vertebral levels (L1-2 through L4-L5), and 'osteophytes' as the presence of osteophytes (grade ≥2) at two or more vertebral levels (L1-2 through L4-L5) [3].

From the 2819 participants of this study, available weight-bearing anterior-posterior radiographs of right/left knees and the pelvis, were scored for knee and hip OA. Radiological knee and hip OA was assessed using the original description of the Kellgren and Lawrence (K&L) grading system [19-21]. Radiographic knee OA was present if the right and/or left knee had a K&L score of ≥ 2. If one of the joints was replaced, the score of the other knee was used in the analyses. The participant was excluded from the analysis if both knees had undergone joint replacement. The same definitions were used for the hip joints. The knee and hip radiographs were scored by several trained readers, who were also blinded to all clinical data of the participants [20, 22].

Pain and morning stiffness

Questions about pain and morning stiffness were asked during an extensive home interview as part of the baseline measurements. The interviewer asked if joint complaints were present during the last months. If the participants answered yes, the interviewer asked whether the pain was present in the following sites: low back, left knee, right knee, left hip, and/or right hip. The participant had to answer the question for each site separately; it was possible to have complaints at several sites. Knee pain or hip pain was positive if pain was present on the left and/or right side. Back pain was positive if the participant had pain in the lower back during the last month.

The interviewer also asked about the presence, duration and location of morning stiffness. If morning stiffness was present, the interviewer asked what its duration was (possible answers were: less than half an hour, half an hour to 1 h or more than 1 h), and where it was located. The location of the stiffness was divided in: 1) legs, 2) arms, 3) back and/or neck, and 4) legs and arms and back. Spinal morning stiffness was present if the partici-

pant answered that the morning stiffness was located at '3' or '4'. Morning stiffness in the legs was defined as stiffness in location '1' or '4'.

Statistical analysis

Multivariable logistic regression analysis was used to explore the associations between morning stiffness and the different radiological features. First, we explored the association between the duration of spinal morning stiffness, a categorical variable, and the two different definitions of LDD (earlier described). Second, we explored the associations between the two definitions of LDD and 1) the presence of spinal morning stiffness, and 2) spinal morning stiffness in combination with low back pain. Third, we assessed whether the association of morning stiffness and LDD was independent of back pain. Finally, we analyzed the association between the two definitions of LDD and morning stiffness lasting < 1 h. Participants with spinal morning stiffness lasting > 1 h were excluded from this analysis. All analyses were adjusted for age, gender, and body mass index (BMI) because earlier studies already reported an association between LDD and these variables [3, 18, 23]. The results of the second and fourth analyses were also presented without adjustment for these variables.

The same four analyses were also used to explore the associations between radiographic knee or hip OA and 1) morning stiffness in the legs, and 2) morning stiffness in the legs in combination with knee or hip pain, respectively.

Results

Population characteristics

Table 1 presents the characteristics of the study participants. The population comprised 1204 men and 1615 women with a mean age of 65.7 years. Low back pain was reported by 499 participants: 173 men and 326 women. Knee pain was reported by 516 participants and hip pain by 328 participants. Spinal morning stiffness was more often present (22.6%) than morning stiffness in the legs (22.0%). When comparing men and women, men showed a higher prevalence of osteophytes (35.6% vs. 25.4%) and a slightly higher percentage of radiographic hip OA (7% vs. 6.7%). Lumbar intervertebral disc space narrowing and radiographic knee OA were more often present in women than in men: 22.5% of the women met the definition of narrowing compared with 13.5% of the men, and 20.1% of the women had radiographic knee OA compared to 10.5% of the men.

Table 1 Characteristics of the study population.

	Men (n =1204) n (%)	Women (n =1615) n (%)	All (n =2819) n (%)
Age: mean ± SD	65.3 ± 6.4	65.9 ± 6.8	65.7 ± 6.6
BMI: mean ± SD*	25.9 ± 2.9	26.6 ± 3.8	26.3 ± 3.5
Pain last month			
Low back pain	173 (14.4)	326 (20.2)	499 (17.7)
Knee pain	154 (12.8)	362 (22.4)	516 (18.3)
Hip pain	84 (7)	244 (15.1)	328 (11.6)
Morning stiffness			
Spinal morning stiffness§	210 (17.4)	426 (26.4)	636 (22.6)
Morning stiffness in legs§	197 (16.4)	424 (26.3)	621 (22.0)
Radiographic features			
'osteophytes'	429 (35.6)	410 (25.4)	839 (29.8)
'narrowing'	162 (13.5)	363 (22.5)	525 (18.6)
Knee K&L ≥2†	126 (10.5)	324 (20.1)	450 (16)
Hip K&L ≥2‡	84 (7)	109 (6.7)	193 (6.8)
Bilateral knee replacement	0 (0)	0 (0)	0 (0)
Bilateral hip replacement	6 (0.5)	18 (1.1)	24 (0.9)

'Osteophytes': the presence of osteophytes (grade ≥2) at two or more vertebral levels on lateral lumbar radiographs.
'Narrowing': disc space narrowing (grade ≥1) at two or more vertebral levels on lateral lumbar radiographs.
* BMI was missing for 12 participants; seven men and five women.
§ Location of morning stiffness was missing for eight participants; four men and four women.
† Knee Kellgren & Lawrence (K&L) score was missing for 169 participants; 58 men and 111 women.
‡ Hip K&L score was missing for seven participants; two men and five women.

LDD and spinal morning stiffness

Table 2 shows the associations between the different durations of spinal morning stiffness and both definitions of LDD. The definition 'narrowing' was more strongly associated than the definition 'osteophytes' for the categories spinal morning stiffness <0.5 h, and spinal morning stiffness ≥0.5 h to ≤1 h. The category spinal morning stiffness > 1 h was more strongly associated with 'osteophytes' than with 'narrowing'.

The associations between the dichotomous variable spinal morning stiffness and both definitions of LDD were statistically significant. The association with 'narrowing' was stronger than the association with 'osteophytes': adjusted odds ratio (aOR) 1.8; 95% confidence interval (CI): 1.4-2.2 and aOR 1.3; 95% CI: 1.1-1.6, respectively. When we also adjusted the analyses

Table 2 Associations regarding different durations of spinal morning stiffness and LDD.

	'Osteophytes'			'Narrowing'		
	Absent n	Present n	aOR (95% CI)	Absent n	Present n	aOR (95% CI)
No spinal morning stiffness	1549	626	Ref. category	1816	359	Ref. category
Spinal morning stiffness lasting < 0.5 h	351	156	1.2 (1.0-1.5)	383	124	1.7 (1.3-2.1)**
Spinal morning stiffness lasting ≥0.5 to ≤ 1 h	55	36	1.7 (1.1-2.7)*	60	31	2.3 (1.4-3.7)**
Spinal morning stiffness lasting >1 h	19	16	2.4 (1.2-4.8)*	26	9	1.9 (0.9-4.3)

'Osteophytes': the presence of osteophytes (grade ≥2) at two or more vertebral levels on lateral lumbar radiographs.
'Narrowing': disc space narrowing (grade ≥1) at two or more vertebral levels on lateral lumbar radiographs.
aOR: adjusted odds ratio; adjusted for age, gender and BMI.
* P<0.05; ** P<0.01.

for back pain, the association became somewhat lower, but stayed statistically significant: aOR1.3; 95% CI: 1.0-1.5 for the definition 'osteophytes' (p-value < 0.05) and aOR1.5;95% CI: 1.2-1.9 for the definition 'narrowing' (p-value < 0.01). The strength of the associations increased when spinal morning stiffness was combined with low back pain: aOR 2.5 95% CI: 1.9-3.4 for 'narrowing' and aOR 1.5; 95% CI: 1.1-2.0 for 'osteophytes'. The association did not increase when analyzing the associations between spinal morning stiffness < 1 h and LDD. All associations are presented in table 3.
The associations decreased when we included only those participants with back pain (n=499) in the analysis: aOR 1.4; 95% CI: 1.0-2.1 for the association between morning stiffness and 'narrowing' and aOR 1.2; 95% CI: 0.8-1.8 for the association between morning stiffness and 'osteophytes'.

Radiological knee and hip OA and morning stiffness in the legs
Table 4 presents data on associations between the different durations of morning stiffness in the legs and radiographic knee and hip OA. The associations between morning stiffness in the legs, and both knee and hip K&L score, were moderate and only statistically significant for knee OA: aOR 1.6; 95% CI: 1.2-2.0 for knee OA, and aOR 1.4; 95% CI: 1.0-1.9 for hip OA. When we also adjusted the analyses for knee/hip pain, the association became somewhat lower and the association between morning

Table 3 Associations between spinal morning stiffness, low back pain and LDD.

	'Osteophytes'				'Narrowing'			
	Absent n	**Present n**	**OR (95% CI)**	**aOR (95% CI)**	**Absent n**	**Present n**	**OR (95% CI)**	**aOR (95% CI)**
No spinal morning stiffness	1549	626	Ref. category	Ref. category	1816	359	Ref. category	Ref. category
Spinal morning stiffness	426	210	1.2 (1.0-1.5)*	1.3 (1.1-1.6)*	472	164	1.8 (1.4-2.2)**	1.8 (1.4-2.2)**
No spinal morning stiffness	1549	626	Ref. category	Ref. category	1816	359	Ref. category	Ref. category
Spinal morning stiffness without low back pain	264	123	1.2 (0.9-1.5)	1.2 (1.0-1.8)	305	82	1.4 (1.0-1.8)*	1.3 (1.0-1.8)*
Spinal morning stiffness in combination with low back pain	162	87	1.3 (1.0-1.8)*	1.5 (1.1-2.0)**	167	82	2.5 (1.9-3.3)**	2.5 (1.9-3.4)**
No spinal morning stiffness	1549	626	Ref. category	Ref. category	1816	359	Ref. category	Ref. category
Spinal morning stiffness lasting < 1 h[§]	406	192	1.2 (1.0-1.4)	1.3 (1.0-1.6)*	443	155	1.8 (2.4-2.2)**	1.8 (1.4-2.2)**
No spinal morning stiffness	1549	626	Ref. category	Ref. category	1816	359	Ref. category	Ref. category
Spinal morning stiffness lasting < 1 h without low back pain	254	117	1.1 (0.9-1.6)	1.2 (1.0-1.5)	292	79	1.4 (1.0-1.8)*	1.4 (1.0-1.8)*
Spinal morning stiffness lasting < 1 h in combination with low back pain[§]	152	75	1.2 (0.9-1.6)	1.4 (1.0-1.9)*	151	76	2.5 (1.9-3.4)**	2.6 (1.9-3.5)**

'Osteophytes': the presence of osteophytes (grade ≥2) at two or more vertebral levels on lateral lumbar radiographs.
'Narrowing': disc space narrowing (grade ≥1) at two or more vertebral levels on lateral lumbar radiographs.
aOR: adjusted odds ratio; adjusted for age, gender and BMI.
* P<0.05; ** P<0.01.
[§] Participants with spinal morning stiffness >1 h are excluded from this analysis.

Table 4 Associations between different durations of morning stiffness in the legs and radiographic knee or hip OA.

	Knee K&L ≥2			Hip K&L ≥2		
	Absent n	**Present n**	**aOR (95% CI)**	**Absent n**	**Present n**	**aOR (95% CI)**
No morning stiffness in the legs	1751	304	Ref. category	2033	139	Ref. category
Morning stiffness in the legs lasting <0.5 h	366	120	1.6 (1.2-2.1)**	460	44	1.4 (1.0-2.0)
Morning stiffness in the legs lasting ≥0.5 ≤1 h	51	22	1.9 (1.1-3.3)*	68	7	1.2 (0.5-2.8)
Morning stiffness in the legs lasting >1 h	24	1	0.2 (0.0-1.6)	23	2	1.4 (0.3-5.9)

aOR: adjusted odds ratio; adjusted for age, gender and body mass index.
*P < 0.05; **P < 0.01.

stiffness in the legs and knee OA was no longer statistically significant: aOR1.2; 95% CI: 0.9-1.6 for radiographic knee OA and aOR1.1; 95% CI: 0.8-1.6 for radiographic hip OA.The strength of the associations increased when morning stiffness in the legs is combined with knee or hip pain. When individuals had both morning stiffness as well as pain in the knee, the association with radiographic knee OA was aOR 3.0; 95% CI: 2.1-4.1. The association between morning stiffness in the legs in combination with hip pain and radiographic hip OA was aOR 3.1; 95% CI: 1.9-5.0. The strength of the associations did not increase much when replacing morning stiffness in the legs with morning stiffness in the legs with a short duration in the analysis (morning stiffness < 0.5 h for the analysis of knee OA and morning stiffness < 1 h for the analysis of knee OA defined according to the ACR criteria [11, 12]). Table 5 presents the associations between morning stiffness in the legs, knee/hip pain and knee/hip K&L score.

Discussion

This study investigated the associations between morning stiffness and different radiological features: LDD, hip K&L score and knee K&L score. We found a moderate association between both definitions of LDD and spinal morning stiffness. The association showed to be independent of back pain, but increased when spinal morning stiffness was combined with low back pain. The definition 'narrowing' was more strongly associated with spinal morning stiffness, and the combination of spinal morning

Table 5 Associations between morning stiffness in the legs, knee pain and radiographic knee OA.

	Knee K&L ≥2			Hip K&L ≥2		
	Absent n	**Present n**	**aOR (95% CI)**	**Absent n**	**Present n**	**aOR (95% CI)**
No morning stiffness in the legs	1751	304	Ref. category	2033	139	Ref. category
Morning stiffness in the legs	444	144	1.6 (1.2-2.0)**	555	53	1.4 (1.0-1.9)
No morning stiffness in the legs	1751	304	Ref. category	2033	139	Ref. category
Morning stiffness in the legs without knee/hip pain	315	64	1.0 (0.7-1.3)	433	28	0.9 (0.6-1.4)
Morning stiffness in the legs in combination with knee/hip pain	129	80	3.0 (2.1-4.1)**	122	25	3.1 (1.9-5.0)**
No morning stiffness in the legs	1751	304	Ref. category	2033	139	Ref. category
Short morning stiffness in the legs §†	366	120	1.6 (1.2-2.1)**	528	51	1.4 (1.0-2.0)
No morning stiffness in the legs	1751	304	Ref. category	2033	139	Ref. category
Morning stiffness in the legs with a short duration, without knee/hip pain	264	52	1.0 (0.7-1.4)	415	27	0.9 (0.6-1.4)
Morning stiffness in the legs with a short duration, in combination with knee/hip pain§†	102	68	3.1 (2.2-4.4)**	113	24	3.2 (2.0-5.3)**

aOR: adjusted for age, gender and BMI.
* P<0.05; ** P<0.01.
§ Morning stiffness in the legs with a short duration, was defined as <0.5 h for the analysis of the knee and <1 h for the analysis of the hip.
† Participants with morning stiffness in the legs >0.5 h were excluded from the analysis of the knee and participants with morning stiffness in the legs >1 h were excluded from the analysis of the hip.

stiffness and low back pain, than was the definition 'osteophytes'. These associations for LDD were similar compared to the associations found for radiographic knee and hip OA. To the best of our knowledge, this is the first study investigating the association between spinal morning stiffness and low back pain with LDD.

Earlier, de Schepper et al. analyzed the association between LDD and low back pain in this same population, reporting an association of odds ratio (OR) 2.1 for the definition 'narrowing' and OR 1.4 for 'osteophytes' [3].

When comparing these associations for back pain with the results of the present study, both associations were higher when spinal morning stiffness was combined with low back pain, compared with the associations for back pain alone. Another study analyzing the association between low back pain and LDD also compared disc space narrowing with the presence of osteophytes [6]. Both these studies found a stronger association between low back pain and LDD for adults with narrowing of the spine than adults with osteophytes [3, 6]. This is consistent with our results, which show a stronger association for 'narrowing' than for 'osteophytes' when analyzing the relation with spinal morning stiffness.

Our results indicated that there is a moderate association between spinal morning stiffness and LDD. This might indicate that spinal morning stiffness is one of the symptoms that clinicians could use for sub-grouping low back pain patients with symptoms due to LDD. However, the association was lower when we only included participants with back pain in the analysis. This might indicate that the presence of morning stiffness is less discriminative in people with back pain. More studies with back pain patients are needed to confirm our association, and to explore whether treatment response or prognosis differs between patients with pain and morning stiffness, and other patients with non-specific low back pain. In this population, a receiver operating characteristic (ROC) curve could be made to examine accuracy of the selection.
Earlier studies of patients with knee pain also reported a similar moderate association between morning stiffness and radiographic knee OA [24, 25]. According to Duncan et al., the relation became stronger when the severity of morning stiffness increased [25]. Reijman et al. also analyzed the associations between different definitions of radiographic hip OA and clinical symptoms, such as pain and morning stiffness, in the Rotterdam Study; they found a moderate association between hip pain and hip K&L score ≥ 2 and a similar association between morning stiffness and hip K&L score ≥ 2 [20]. We expected to find a difference between the associations of (1) morning stiffness, and (2) morning stiffness with a short duration, with the radiographic features LDD, knee or hip OA, because morning stiffness in the knee < 0.5 h or hip < 1 h is an ACR criterion for knee or hip OA [11, 12] and spinal morning stiffness > 1 h is one of the criteria for ankylosing spondylitis [26, 27] however, no such a difference was found. It must be noted, however, that power for this stratified analysis was limited, and so no final conclusion can be drawn from this result.

Limitations

Our study had a few limitations which might influence the results. First, only lateral radiographs of the lumbar spine were assessed. Therefore, single-sided disc space narrowing and lateral osteophytes may have been missed. Second, because only lateral radiographs of the lumbar spine were available, we could not score the facet joints, which are the only synovial joints in the spine. Therefore we could not examine if the presence of facet joint OA is responsible for the association between spinal morning stiffness and LDD or whether LDD is associated with morning stiffness independently of facet joint OA.

A third limitation is that, for another study purpose, only baseline radiographs of participants with baseline and 6.6 years follow-up measurements were scored. On average, participants who were available for 6.6-years follow-up measurements were younger and healthier during baseline than those participants who were not available for follow-up measurements. This caused some selection bias in our study sample.

The fourth limitation is that the location of morning stiffness was described as 'spinal morning stiffness' and 'morning stiffness in the legs' without distinction between the precise locations. Therefore we are unable to differentiate between morning stiffness in the hip/knee, or morning stiffness in the cervical, thoracic or lumbar spine. When we analyzed the association between morning stiffness in the legs and radiographic OA in the lower body (hip and/or knee OA), it did not result in a much higher association: aOR 1.6; 95% CI: 1.3-2.0. Another limitation related to the location of morning stiffness is that participants who indicated that the morning stiffness was located in the arms, spine, and the legs (location 4) had a positive score for both spinal morning stiffness and morning stiffness in the legs. If we had more precise information about the location of the morning stiffness and radiographic information of the facet joint, the associations might have been different.

In conclusion, spinal morning stiffness is frequently reported in this study population. According to our analyses, there appears to be a small association between spinal morning stiffness and LDD. The magnitude of the association was higher when spinal morning stiffness was combined with low back pain.

References

1. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 2003;102: 167-78.
2. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* 2006;332:1430-4.
3. de Schepper EIT, Damen J, van Meurs JBJ, Ginai AZ, Popham M, Hofman A, et al. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. *Spine* 2010;35: 531-6.
4. Hassett G, Hart DJ, Manek NJ, Doyle DV, Spector TD. Risk factors for progression of lumbar spine disc degeneration: the Chingford Study. *Arthritis Rheum* 2003;48:3112-7.
5. Hicks GE, Morone N, Weiner DK. Degenerative lumbar disc and facet disease in older adults: prevalence and clinical correlates. *Spine* 2009;34:1301-6.
6. Pye SR, Reid DM, Smith R, Adams JE, Nelson K, Silman AJ, et al. Radiographic features of lumbar disc degeneration and self-reported back pain. *J Rheumatol* 2004;31:753-8.
7. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine* 1997;22:427-34.
8. Eubanks JD, Lee MJ, Cassinelli E, Ahn NU. Does lumbar facet arthrosis precede disc degeneration? A postmortem study. *Clin Orthop Relat Res* 2007; 464:184-9.
9. Fujiwara A, Tamai K, Yamato M, An HS, Yoshida H, Saotome K, et al. The relationship between facet joint osteoarthritis and disc degeneration of the lumbar spine: an MRI study. *Eur Spine J* 1999;8:396-401.
10. Suri P, Miyakoshi A, Hunter DJ, Jarvik JG, Rainville J, Guermazi A, et al. Does lumbar spinal degeneration begin with the anterior structures? A study of the observed epidemiology in a community-based population. *BMC Musculoskelet Disord* 2011;12:202.
11. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;34:505-14.
12. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;29:1039-49.

13. Hofman A, Breteler MM, van Duijn CM, Krestin GP, Pols HA, Stricker BH, et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007; 22:819-29.
14. Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HL, Klaver CC, et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* 2011;26:657-86.
15. Lane NE, Nevitt MC, Genant HK, Hochberg MC. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol* 1993;20:1911-8.
16. Kettler A, Wilke HJ. Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. *Eur Spine J* 2006;15:705-18.
17. Cohn EL, Maurer EJ, Keats TE, Dussault RG, Kaplan PA. Plain film evaluation of degenerative disk disease at the lumbosacral junction. *Skeletal Radiol* 1997; 26:161-6.
18. Symmons DP, van Hemert AM, Vandenbroucke JP, Valkenburg HA. A longitudinal study of back pain and radiological changes in the lumbar spines of middle aged women. II. Radiographic findings. *Ann Rheum Dis* 1991;50: 162-6.
19. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16:494-502.
20. Reijman M, Hazes JM, Pols HA, Bernsen RM, Koes BW, Bierma-Zeinstra SM. Validity and reliability of three definitions of hip osteoarthritis: cross sectional and longitudinal approach. *Ann Rheum Dis* 2004;63:1427-33.
21. Schiphof D, de Klerk BM, Kerkhof HJ, Hofman A, Koes BW, Boers M, et al. Impact of different descriptions of the Kellgren and Lawrence classification criteria on the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2011;70: 1422-7.
22. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. *Ann Rheum Dis* 1998;57:203-8.
23. Pye SR, Reid DM, Adams JE, Silman AJ, O'Neill TW. Influence of weight, body mass index and lifestyle factors on radiographic features of lumbar disc degeneration. *Ann Rheum Dis* 2007;66:426-7.
24. Claessens AA, Schouten JS, van den Ouweland FA, Valkenburg HA. Do clinical findings associate with radiographic osteoarthritis of the knee? *Ann Rheum Dis* 1990;49:771-4.
25. Duncan R, Peat G, Thomas E, Hay E, McCall I, Croft P. Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? *Ann Rheum Dis* 2007;66:86-91.

26. Chavannes AW, Mens JMA, Koes BW, Lubbers WJ, Ostelo R, Spinnewijn WEM, et al. NHG-Standaard Aspecifieke lagerugpijn; guideline for non-specific low back pain of the Dutch College of General Practitioners. *Huisarts Wet* 2005;48:113-23.
27. Goei The HS, Steven MM, van der Linden SM, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a comparison of the Rome, New York and modified New York criteria in patients with a positive clinical history screening test for ankylosing spondylitis. *Br J Rheumatol* 1985;24:242-9.

8

Disc degeneration of the upper lumbar discs is associated with hip pain

Evelien I.T. de Schepper, Jurgen Damen,
Pieter K. Bos, Albert Hofman,
Bart W. Koes, Sita M. Bierma-Zeinstra

Abstract

Purpose

A possible cause of hip pain is the presence of radiating pain from the higher lumbar spine. Identification of factors associated with hip pain arising from the lumbar spine would aid the physician. The first step in identifying possible factors is to look at the association between hip pain and osteoarthritis of the lumbar spine.

Methods

In an open population based study of people 55 years and older (Rotterdam study), 2819 lumbar radiographs were scored for the presence and severity of individual radiographic features of disc degeneration. Hip osteoarthritis was scored on anteroposterior pelvic radiographs, and questionnaires including self-reported hip pain were taken. Logistic regression adjusted for possible confounders was used to determine the association between self-reported hip pain and the individual radiographic features of lumbar disc degeneration.

Results

The presence of disc space narrowing grade ≥ 1 at level L1/L2 was significantly associated with hip pain in the last month (men OR = 2.0; 95% CI: 1.1 to 3.8 and women OR = 1.7; 95% CI: 1.1 to 2.5). The presence of disc space narrowing grade ≥ 1 at level L2/L3 was only significantly associated with hip pain in women. The strength of the associations increased for self-reported chronic hip pain, especially in men (L1/L2 OR = 2.5; 95% CI: 1.3 to 5.0). The presence of disc space narrowing at the lower levels (L3/L4/L5/S1) was not significantly associated with hip pain.

Conclusion

Our data provide evidence for an association between hip pain and disc space narrowing at disc level L1/L2 and L2/L3. In case of uncertainty of the cause of hip pain, evaluation of lumbar radiographs may help to identify those hip pain patients who might have pain arising from the lumbar spine.

Introduction

Hip pain is a common symptom among older adults, with a point prevalence of 14.3% reported in the United States [1]. The differential diagnosis of hip pain is broad and includes intra-articular pathology, extra-articular pathology and other causes like radiating pain from the lumbar spine. Differentiating back pain from hip pain in patients who present with classic signs and symptoms is mostly not difficult and generally does not require further testing to establish an accurate diagnosis. However, in some cases, patients present with nonspecific complaints of pain in the lumbar spine, buttock, lateral hip, or thigh [2]. The differentiation of signs and symptoms suggestive of hip disorders versus spine disorders is important in giving patients the most beneficial treatment, especially if the treatment includes a major reconstructive surgery, such as hip replacement.

Differentiating whether hip-pain originates from the hip, the spine or both may be challenging. Brown et al. [3] attempted to determine which physical signs and symptoms best predict the primary source of pain in patients with hip-, spine- or concomitant disorders. After final diagnosis with imaging studies, they found that although limited internal rotation, groin pain and a limp are more commonly associated with a hip disorder, these symptoms are also seen in patients with spine alone or both hip- and spine-disorders.

To make a differentiation between hip and spine originated hip pain there have been a few studies about the usefulness of local anaesthetic with(out) corticosteroid hip infiltrations, to differentiate intra-articular causes of hip pain from spinal causes [4,5,6,7]. To our knowledge, there have been no studies about the usefulness of local spine infiltrations to differentiate hip and spine originated hip pain. However, infiltration of every patient with atypical hip pain for possible coexistent lumbar spine osteoarthritis would be counterproductive and costly. Preoperative identification of factors associated with hip pain arising from the lumbar spine would aid the physician by identifying the subgroup of patients who might not experience full relief of pain with a hip arthroplasty.

One of the first steps to identify possible factors is to look at the association between hip pain and osteoarthritis of the lumbar spine. The purpose of this study was to explore the association of self-reported hip pain with the different individual radiographic features (IRF) of spinal osteoarthritis by vertebral level, including osteophytes and disc space narrowing.

Materials and methods

Study population

The data for this study originate from data of the Rotterdam Study, an open population prospective cohort of people aged 55 years and older. The objective of the Rotterdam Study is to investigate the incidence of, and risk factors for, chronic disabling diseases. The study design has been described previously [8]. All 10,275 inhabitants of Ommoord (a district in Rotterdam, the Netherlands) were invited to participate. The baseline measurements were conducted between 1990 and 1993. In total, 7983 participants were examined.

For this study, 2819 lumbar radiographs were scored. The selection was based on the availability of radiographs of the hip and spine at a follow-up measurement 6.6 years later [9,12].

Radiographic scoring

Lumbar lateral radiographs were scored by a single observer trained by a radiologist for the presence of the individual radiographic features of disc degeneration. The observer was blinded to clinical characteristics of the participants. Each vertebral level from L1/2 to L5/S1 was reviewed for the presence and severity of osteophytes (anterior) and vertebral narrowing, using the Lane atlas [10,11]. In this atlas grade 0 = none; grade 1 = mild; grade 2 = moderate; and grade 3 = severe. The lumbosacral disc space was defined as narrowed when its height was less than that of the disc space between the third and fourth lumbar vertebrae. This is due to a normal progression of increasing disc-space height from the third and fourth to the fourth and fifth lumbar vertebrae, and then a relative narrowing of the height of the lumbosacral disc space. Sclerosis was not scored because of the earlier reported low ICC for this feature [11].

Inter-observer reproducibility was assessed by a second independent observer who evaluated a random selection of 140 (5%) X-rays. The ICC was 0.83 for osteophytes and 0.77 for vertebral narrowing, indicating good reproducibility.

Weight bearing anteroposterior radiographs of the pelvis were obtained. One trained reader evaluated the radiographs obtained at baseline, unaware of the clinical status of the participants [9]. At baseline, radiological osteoarthritis of the hip was quantified by measurements following the Kellgren & Lawrence grading system (atlas-based) in five grades (from 0 to

4). A person was considered to have osteoarthritis of the hip, if the Kellgren & Lawrence score of one or both joints was equal to or larger than two [9].

Hip pain

Hip pain and low back pain were determined from interviewing the participants during the home visits. Participants were asked "Did you have complaints of the (left and/or right) hip during the last month?". Hip pain was defined to be present if the answer was positive. Participants were subsequently asked "What is the duration of the present hip complaints?". For low back pain similar questions were asked. We defined chronic hip pain to be present if the duration of the hip joint pain was more than one year.

Participants also visited the research center, where X-rays were obtained. Height and weight were measured with participants wearing indoor clothing and without shoes. Body mass index (BMI) was calculated as weight in kilograms divided by length in meters squared (kg/m^2).

Statistical Analyses

We defined disc space narrowing to be present if the grade was mild, moderate, or severe (grade ≥1). Because of the small proportion of subjects without osteophytes, we used a higher cutoff value for this feature. We defined osteophytes to be present if the grade was moderate or severe (grade ≥2) [12]. Using these definitions we calculated the prevalence of the IRF by vertebral level (L1/2 to L5/S1) and gender.

In order to explore the association between the IRF by vertebral level and hip pain, hip pain was used as the dependent variable with adjustments made for age and gender. The assessments of the associations were also adjusted for radiological hip osteoarthritis, as this variable was shown to be associated with disc space narrowing and of course with hip pain. The same was true for low back pain [12]. In addition, the assessments of the association were also adjusted for BMI, as this variable has been reported to be associated with both hip pain and some of the individual radiographic features [13,14,15].

In a separate analysis we explored the association between the IRF by vertebral level and hip pain in subjects with no sign of radiological hip osteoarthritis. The results of the analyses are expressed as odds ratios (OR) with 95% confidence intervals (CI), stratified for gender. The same methods were used to explore the association between the IRF and chronic

hip pain. Statistical analysis was performed using SPSS version 15 (SPSS Inc, Chicago, USA).

Results

Subject characteristics

Baseline characteristics are shown in Table 1. There were 1204 men (mean age 65.3 years, standard deviation (SD) 6.4) and 1615 women (mean age 65.9 years, SD 6.8). Hip pain during the last month was reported more often by women than men (244 (15.1%) vs. 84 (7.0%) $p < 0.05$) (Table 1). Chronic hip pain was reported in the majority (82%) of the current hip pain cases and was also more often reported by women (208 (12.9%) vs. 62 (5.1%) $p < 0.05$). Radiological hip osteoarthritis was observed in 209 (7.4%) persons (Kellgren & Lawrence ≥2 in one or both hips).

Influence of gender and vertebral level

The prevalence of the IRF in men and women is shown in Table 1. Osteophytes were the most frequent observed radiographic feature and were slightly more common in men than women (95% vs. 91%; $p<0.05$). Disc space narrowing was more frequent in women than men (65% vs. 53%; $p<0.05$). In terms of their distribution by vertebral level, narrowing was more frequent at the lower lumbar disc levels.

Disc space narrowing grade ≥ 1 at level L1/L2 was more common in persons with hip pain (19% vs. 10%; $p <0.05$). And hip pain was more common in persons with disc space narrowing grade ≥ 1 at level L1/L2 (21% vs. 11%; $p <0.05$).

Association with LDD

Table 2 shows the association between hip pain and the IRF, adjusted for age, gender, BMI, hip arthritis and low back pain. The presence of disc space narrowing grade ≥ 1 at level L1/L2 was significantly associated with hip pain in the last month, both in men and women (men OR = 2.0; 95% CI: 1.1 to 3.8 and women OR = 1.7; 95% CI: 1.1 to 2.5) (Table 2). The presence of disc space narrowing grade ≥ 1 at level L2/L3 was significantly associated with hip pain in the last month, only in women (OR = 1.6; 95% CI: 1.1 to 2.2). The strength of the associations increased for the participants with chronic hip pain, especially for men (L1/L2 OR = 2.5; 95% CI: 1.3 to 5.0). The strength of the associations also increased for the group of subjects with no radiological hip osteoarthritis (men chronic pain L1/L2

Table 1. Frequency of hip pain and individual radiographic features of the low back in men and women

	Men, N = 1204	**Women, N = 1615**	**All, N = 2819**	**Hip pain, N = 328**
Age (years) Mean ± SD	65.3 ± 6.4	65.9 ± 6.8	65.7 ± 6.6	66.2 ± 6.8
Body mass index (BMI) Mean ± SD	25.9 ± 2.9	26.6 ± 3.8	26.3 ± 3.5	27.0 ± 3.9
Hip pain (%)†	84 (7.0)	244 (15.1)	328 (11.6)	328 (100)
Chronic hip pain (%)‡	62 (5.1)	208 (12.9)	270 (9.6)	270 (82.3)
Hip osteoarthritis (%)	94 (7.8)	115 (7.1)	209 (7.4)	51 (15.5)
Osteophytes low back (%)				
Grade ≥ 1	1148 (95.3)	1467 (90.8)	2615 (92.8)	306 (93.3)
Grade ≥ 2	832 (69.1)	929 (57.5)	1761 (62.5)	217 (66.2)
Grade 3	536 (44.5)	505 (31.3)	1041 (36.9)	134 (40.9)
Narrowing low back (%)				
Grade ≥ 1	637 (52.9)	1048 (64.9)	1685 (59.8)	210 (64.0)
Grade ≥ 2	286 (23.8)	525 (32.5)	811 (28.8)	115 (35.1)
Grade 3	40 (3.3)	107 (6.6)	147 (5.2)	20 (6.1)
Osteophytes ≥2 (%)				
L1-2	282 (23.4)	297 (18.4)	579 (20.5)	84 (25.6)
L2-3	347 (28.8)	404 (25.0)	751 (26.6)	105 (32.0)
L3-4	428 (35.5)	364 (22.6)	792 (28.1)	100 (30.5)
L4-5	403 (33.5)	354 (21.9)	757 (26.9)	94 (28.7)
L5-S1	312 (25.9)	303 (18.8)	615 (21.8)	68 (20.7)
Narrowing ≥1 (%)				
L1-2	107 (8.9)	201 (12.5)	308 (10.9)	63 (19.2)
L2-3	135 (11.3)	307 (19.0)	442 (15.7)	81 (24.7)
L3-4	153 (12.7)	342 (21.1)	495 (17.6)	78 (23.8)
L4-5	268 (22.2)	526 (32.6)	794 (28.2)	111 (33.8)
L5-S1	408 (34.0)	662 (41.0)	1070 (38.0)	127 (38.7)

† Hip pain: complaints of the hip joint during last month
‡ Chronic hip pain: duration present hip joint complaints > 1 year

OR = 2.7; 95% CI; 1.3 to 5.5 and women chronic pain L1/L2 OR = 2.0; 95% CI; 1.3 to 3.2).

The presence of disc space narrowing at the lower levels (L3/L4/L5/S1) was not significantly associated with hip pain. The presence of disc space narrowing grade ≥ 2 was not explored, because of the low number of persons with disc space narrowing grade ≥ 2 at the upper levels.

The presence of osteophytes grade ≥ 2 was not significantly associated with hip pain at any level (data not shown).

Table 2 Association between disc space narrowing and hip pain

	Men, N = 1204		
Narrowing level	N (%)	Hip pain OR (95% CI)	Chronic hip pain OR (95% CI)
L1-L2	107 (8.9)	2.0 (1.1 – 3.8)*	2.5 (1.3 – 5.0)**
L2-L3	135 (11.3)	0.9 (0.4 – 1.8)	1.1 (0.5 – 2.4)
L3-L4	153 (12.7)	1.1 (0.6 – 2.1)	1.1 (0.5 – 2.2)
L4-L5	268 (22.2)	1.2 (0.7 – 2.0)	1.4 (0.8 – 2.5)
L5-S1	408 (33.9)	0.7 (0.4 – 1.1)	0.6 (0.4 – 1.1)
	Women, N = 1615		
Narrowing level	N (%)	Hip pain OR (95% CI)	Chronic hip pain OR (95% CI)
L1-L2	201 (12.5)	1.7 (1.1 – 2.5)*	1.8 (1.1 – 2.7)**
L2-L3	307 (19.0)	1.6 (1.1 – 2.2)*	1.6 (1.1 – 2.3)*
L3-L4	342 (21.1)	1.0 (0.7 – 1.4)	1.1 (0.7 – 1.5)
L4-L5	526 (32.6)	0.9 (0.7 – 1.3)	1.0 (0.7 – 1.4)
L5-S1	662 (41.0)	1.0 (0.7 – 1.3)	0.9 (0.7 – 1.2)
	All, N = 2819		
Narrowing level	N (%)	Hip pain OR (95% CI)	Chronic hip pain OR (95% CI)
L1-L2	308 (10.9)	1.8 (1.3 – 2.5)**	2.0 (1.4 – 2.8)**
L2-L3	442 (15.7)	1.4 (1.0 – 1.9)*	1.5 (1.1 – 2.1)*
L3-L4	495 (17.6)	1.1 (0.8 – 1.4)	1.1 (0.8 – 1.5)
L4-L5	794 (28.2)	1.0 (0.8 – 1.3)	1.1 (0.8 – 1.5)
L5-S1	1070 (38.0)	0.9 (0.7 – 1.1)	0.8 (0.6 – 1.1)

Adjusted for age, gender, BMI, hip arthrosis and low back pain
* p = < 0.05
** p = < 0.01
OR odds ratio
CI confidence interval

Discussion

The differentiation of signs and symptoms suggestive of hip disorders versus spine disorders is important in giving patients the most beneficial treatment. The purpose of this study was to explore the association of self-reported hip pain with the different individual radiographic features (IRF) of spinal osteoarthritis. In this study, disc space narrowing at level L1/L2 appeared to be associated with pain in the hip region, especially in men. The strength of the associations increased for participants with chronic hip pain and in those without radiological signs of hip osteoarthritis. These results suggest that in case of uncertainty of the cause of hip pain, evaluation of lumbar radiographs may help to identify those hip pain patients that may benefit most from further diagnostic evaluation.

Our data provides evidence for radiating pain from the higher lumbar spine as a possible cause of hip pain in a cross-sectional open population based study. One of the explanations that can be found for the association between hip pain and disc space narrowing at level L1/L2 and L2/L3 is "referred pain". The term "referred pain" is used for pain localized not in the site of its origin but in areas that may be adjacent or at a distance from such a site. Several theories have been proposed to explain the "referred pain" phenomenon, with the convergence-projection theory the most widespread [16,17]. Input from different tissue types converge on the same dorsal horn neurons [18]. And after activation, increased nociceptive input is transmitted supraspinally and misinterpreted at the cortical level as pain from other tissues. It is possible that the reduction of space between the vertebrae as a consequence of the degenerative disc leads to increased pressure on spinal ligaments and other supporting tissues. This can be misinterpreted at the cortical level as pain from other tissues, like the hip region. Experimental studies have confirmed that noxious stimulation of interspinous ligament, facet joint, and paravertebral muscles causes referred pain that can radiate into the extremity [19,20,21].

Another explanation for the radiating pain from the higher lumbar spine can be found in the dermatomal innervations of the hip region. It is suggested that impingement of the higher lumbar spinal nerve roots (L1-L3) can cause pain in the dermatomal distribution surrounding the hip. The dermatomal distribution of the L1 spinal nerve is located in the groin and the upper part of the buttock. The distribution of the L2 spinal nerve is located in the outside thigh. It is possible that reduction of space between the vertebrae as a consequence of the degenerative disc is more likely to lead to impingement of the L1 and L2 nerve roots, and therefore causes pain in the dermatomal distribution. Spinal nerve roots pass through the intervertebral foramen as they travel from the spinal cord toward the periphery. It has been reported that narrowing of the disc space can reduce the vertical diameter of this intervertebral foramen [22].

The explanation for the stronger association between hip pain and disc space narrowing compared with the presence of osteophytes is unknown. This study evaluates the severity of anterior osteophytes, unfortunately we could not evaluate any bony aspects of the intervertebral foramen. The explanation for the stronger association between hip pain and disc space narrowing at L1/L2 in men compared with the association in women is

also unknown. It is possible that even though women reported hip pain more often, only a small proportion of the complaints are due to disc space narrowing, whereas other factors determine the feeling of pain. Men and women could also report pain differently therefore effecting the association between hip pain, disc space narrowing and gender. Cecchi et al., showed that women presented with significantly more severe pain than men [23]. Finally, the explanation for the absence of an association between hip pain and disc space narrowing at L2/L3 in men compared to women is also unknown. It is maybe due to an evidently lower prevalence of disc space narrowing at L2/L3 in men compared to women.

Our study had several advantages. It was population based with a relatively high number of subjects. We used a semi-quantitative score, using standard radiographs, to characterize the presence and severity of hip and spine osteoarthritis. Assessment of the radiographs was carried out without knowledge of the questionnaire data, and so any errors in classification are likely to have been non-directional. We defined chronic hip pain and chronic low back pain to be present if the duration of the hip joint pain was more than one year. In literature, others have chosen three months or even six months as the dividing line between acute and chronic pain [24]. However, with our definition, chronic pain included long lasting chronic complaints with long lasting impact on ones life.

However, there are several limitations in our explorative study that need to be considered when interpreting the results. Our data did not include the precise location of the hip pain. This limitation is partly undermined by the fact that the dermatomal distribution of L1 and L2 includes the upper part of the buttock, the groin and the lateral thigh, which covers a wide area of the hip region. Further, our data did not include a clinical evaluation of the hip pain. In this way we could not account for the potential of soft-tissue pathology contributing to the reported hip pain.
Moreover, hip osteoarthritis was only considered when the Kellgren & Lawrence score of one or both joints was equal to or larger than two in agreement with conventional epidemiological definitions for hip osteoarthritis [25]. In this way there is still a possibility of the presence of hip osteoarthritis which is not clearly visible yet on radiographs at that time point. To exclude the possibility of this confounding, we reanalyzed the data with adjusting for presence of radiographic hip osteoarthritis 6.6 years later. We defined a new variable that included all the participants with hip

osteoarthritis at baseline and/or hip osteoarthritis 6.6 years later (n = 413). The strength of the associations was unchanged (for chronic pain the L1/L2 OR was 1.9; 95% CI; 1.3 to 2.7; again higher in men (OR = 2.7; 95% CI; 1.4 to 5.3) than in women (OR = 1.7; 95% CI; 1.1 to 2.6).
Furthermore, there could be some selection bias in favor of relatively healthy participants. The participants in the present study had to be mobile enough to visit the research center for X-ray examination, both for the baseline and follow up appointments (mean 6.6 years) [9]. In other words, patients with the most severe symptoms were most likely not included, but this may be inevitable in long-term prospective cohort studies.

What are the implications of these findings for researchers and clinicians?
Accurate diagnosis of pain originating from the hip joint can be clinically challenging. There have been several studies about the usefulness of hip injections to differentiate intra-articular causes of hip pain from spinal causes [4,5,6,7]. To our knowledge, there have been no studies about the usefulness of local spine infiltrations to differentiate hip and spine originated hip pain. A possible explanation for this is the availability of a successful treatment for degenerated hip disease (hip arthroplasty), but less predictable treatment options for degenerative spine disorders.
The differentiation of signs and symptoms suggestive of a hip disorder is important in giving patients adequate information regarding their condition and for applying the most beneficial treatment. Our data provides evidence for an association between hip pain and disc space narrowing at disc level L1/L2 and L2/L3. In case of uncertainty of the cause of hip pain, evaluation of lumbar radiographs may help to identify those hip pain patients who might have pain arising from the lumbar spine. Perhaps hip infiltration in patients without higher lumbar disc degeneration is even unnecessary. However, well designed studies are needed to verify this hypothesis.

Conclusion
In conclusion, this study explores the association of self-reported hip pain with lumbar spine osteoarthritis. Our data provides evidence for an association between hip pain and disc space narrowing at disc level L1/L2 and L2/L3. In case of uncertainty of the cause of hip pain, evaluation of lumbar radiographs may help to identify those hip pain patients who might have pain arising from the lumbar spine. Well designed studies are needed to verify this hypothesis.

References

1. Christmas C, Crespo CJ, Franckowiak SC, et al. How common is hip pain among older adults? Results from the Third National Health and Nutrition Examination Survey. *J Fam Pract* 2002;51(4):345-348.
2. Illgren R, Honkamp N, Weisman M, et al. The Diagnostic and Predictive Value of Hip Anesthetic Arthrograms in Selected Patients Before Total Hip Arthroplasty. *The Journal of Arthroplasty*. 2006;21(5):724-730
3. Brown MD, Gomez-Marin O, Brookfield KF, et al. Differential diagnosis of hip disease versus spine disease. *Clin Orthop Relat Res* 2004(419):280-284.
4. Ashok N, Sivan M, Tafazal S. The diagnostic value of anaesthetic hip injection in differentiating between hip and spinal pain. *Eur J Orthop Surg Traumatol*. 2009;19:167-171
5. Faraj A, Kumaraguru P, Kosygan K. Intra-articular bupivacaine hip injection in differentiation of coxarthrosis from referred thigh pain: A 10 year study. *Acta Orthopaedica Belgica*. 2003;69(6):518-521
6. Kleiner J, Thorne R, Curd J. The Value of Buvicaine Hip Injection in the Differentiation of Coxarthrosis from Lower Extrimity Neuropathy. *The Journal of Rheumatology*. 1991;18(3): 422-427
7. Pateder D, Hungerford M. Use of Fluoroscopically Guided Intra-articular Hip Injection in Differentiating the Pain Source in Concomitant Hip and Lumbar Spine Arthritis. *Am J Orthop*. 2007;36(11):591-593
8. Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007;22(11):819-829.
9. Reijman M, Hazes JM, Pols HA, et al. Role of radiography in predicting progression of osteoarthritis of the hip: prospective cohort study. *Bmj* 2005; 330(7501):1183.
10. Kettler A, Wilke HJ. Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. *Eur Spine J* 2006;15(6):705-718.
11. Lane NE, Nevitt MC, Genant HK, et al. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol* 1993;20(11):1911-1918.
12. de Schepper E, Damen J, van Meurs JB, et al. The association between lumbar disc degeneration and low back pain: the influence of age, gender and individual radiographic features. *Spine* 2010;35(5):531-6.
13. Lievense AM, Bierma-Zeinstra SM, Verhagen AP, et al. Influence of obesity on the development of osteoarthritis of the hip: a systematic review. *Rheumatology (Oxford)* 2002;41(10):1155-1162.

14. Pye SR, Reid DM, Adams JE, et al. Influence of weight, body mass index and lifestyle factors on radiographic features of lumbar disc degeneration. *Ann Rheum Dis* 2007;66(3):426-427.
15. Tukker A, Visscher T, Picavet H. Overweight and health problems of the lower extremities: osteoarthritis, pain and disability. *Public Health Nutr* 2008:1-10.
16. Frey Law LA, Sluka KA, McMullen T, et al. Acidic buffer induced muscle pain evokes referred pain and mechanical hyperalgesia in humans. *Pain*. 2008 Nov 30;140(2):254-64
17. Graven-Nielsen T. Fundamentals of muscle pain, referred pain, and deep tissue hyperalgesie. *Scand J Rheumatol* 2006;35:1-43
18. Willis WD, Coggeshall RE. *Sensory mechanisms of the spinal cord*. New York: Kluwer Academic/Plenum Publishers; 2004.
19. Feinstein R, Langton J, Jameson R, et al. Experiments on pain referred from deep somatic tissues. *J Bone Joint Surg Am* 1954;36A:981-97
20. Kellgren J. On the distribution of pain arising from deep somatic structures with charts of segmental pain areas. *Clin Sci* 1939;4:35-46
21. O'Neill CW, Kurgansky ME, Derby R, et al. Disc stimulation and patterns of referred pain. *Spine*. 2002;27:2776-81
22. Cinotti G, De Santis P, Nofroni I, et al. Stenosis of lumbar intervertebral foramen. *Spine* 2002;27(3):223-229
23. Cecchi F, Debolini P, Lova RM, et al. Epidemiology of back pain in a representative cohort of Italian persons 65 years of age and older: the InCHIANTI study. *Spine* 2006;31(10):1149-1155.
24. Bonica J. *The management of pain*. Philadelphia: Lea and Febiger, 1990.
25. Reijman M, Hazes JM, Pols HA, et al. Validity and reliability of three definitions of hip osteoarthritis: cross sectional and longitudinal approach. *Ann Rheum Dis*. 2004 Nov;63(11):1427-33

9 General Discussion

The aim of the work presented in this thesis was to gain more insight into the diagnostic tools used in patients with low back pain. First, we examined how lumbar MRI is currently used by general practitioners in the Netherlands. Then, we updated our earlier systematic review on the diagnostic accuracy of tests used to diagnose lumbar spinal stenosis. And finally, we performed an epidemiological study to investigate the characteristics of lumbar disc degeneration. The previous chapters report on the findings of each study that was conducted to achieve this objective. This chapter discusses how to interpret these results in the context of existing literature and in light of some important methodological issues. Finally, we present some implications for future research and clinical practice.

Background

The general practitioner is the healthcare provider that is most often visited for back pain [1-3]. Low back pain is usually divided into 'specific' and 'non-specific' low back pain. Specific low back pain is defined as symptoms caused by a specific pathophysiological mechanism, such as a hernia nuclei pulposi, discitis, malignancy, or fracture. The most common specific cause of low back pain is a herniated disc.
Overall, adequate treatment of patients with low back pain begins with making the correct diagnosis. The diagnosis determines the appropriate steps to be taken, including reassurance and advice, but also referral to a second line treatment or to additional diagnostics. How can we make a distinction between the small group of patients with a specific disorder and the large group of patients with non-specific low back pain?
In the Netherlands the first distinction is made by the general practitioner, almost exclusively based on history and physical examination. In addition, in a small proportion of patients, diagnostic imaging is needed. International guidelines recommend the use of imaging only when there is suspicion of serious pathology (fracture, malignancy and discitis), or in patients with severe sciatica for whom surgery is indicated because they fail to respond to conservative care for at least 6-8 weeks [4, 5]. At the moment, there is widespread consensus that there is no indication to perform diagnostic imaging in patients with non-specific back pain.

MRI in primary care

If a physician suspects the presence of a specific disease, diagnostic imaging can be used. In recent years substantial improvements have been made in the techniques used to visualize the anatomy of the lower spine, and these techniques are now widely available. Advanced imaging with MRI is widely used by clinicians from all over the world and also general practitioners in the Netherlands are nowadays able to refer patients for MRI of the lumbar spine. However, data on the use of MRI by these general practitioners was still lacking. Therefore, in the first part of this thesis, we examined how lumbar MRI is currently used by general practitioners in the Netherlands. For this, we designed an observational prospective cohort study with a 12-month follow-up in low back pain patients referred for MRI in general practice.
One of the first, most striking, results was that in our cohort of patients referred for lumbar MRI by their general practitioner, most of them reported long-lasting, severe low back complaints **(Chapter 2)**. When we compare these results with the results of other back pain cohort studies, there are some noteworthy differences (Table 1). In our cohort, the baseline back pain severity and disability scores were generally higher than reported in earlier cohort studies on low back pain [6-10]. Also, compared with the study by Jarvik et al. [9], our patients more often reported pain radiating to the leg below the knee and more often had a history of back surgery.

Furthermore, the MRI reports of our cohort showed a relatively high number of disc pathologies, i.e. the MRI reports showed disc herniation in 72% and nerve root compression in 69% of the patients. These prevalences are much higher than reported in persons without low back pain [11] (Table 2). They are also higher than reported in cohorts of patients with acute back pain [12], and in cohorts including primary care patients referred for radiography [9, 13]. However, our prevalences did match those reported by Hancock et al. [14] who compared rates of MRI findings of 30 patients with low back pain and 30 pain-free controls.
These results show that a large proportion of the referred patients in our cohort had complaints for which imaging could be indicated. Specifically: 374 patients (55%) reported sciatica complaints for at least 6 weeks and could therefore be candidates for surgery, in which case imaging is recommended. These results are an indication that general practitioners in the Netherlands tend to adhere to their guidelines. However, to validate this

Table 1 Baseline characteristics of patients in low back pain cohorts

	de Schepper et al. [Chapter 2] n=683	Campbell et al. [9] N=488	Hill et al. [10] N=2526	Mehling et al. [11] N=605	Jarvik et al. [12] N=380	Jarvik et al. [13] N=349
Study characteristics	Referred by general practitioner for MRI	LBP in primary care	LBP in primary care	Acute LBP in primary care	Referred by physician for radiography	Older adults referred for early MRI/CT
Patient characteristics						
Age in years, mean (SD)	49.9 (12.5)	47.4 (9.0)	44.5 (9.7)	50.5 (13)	53.2 (15)	72.8 (6.0)
Male	365 (53)	185 (38)	1067 (42)	266 (44)	169 (44)	120 (34)
Body mass index, mean (SD)	25.9 (3.8)	.	.	.	28.7 (6.0)	.
Education level low	244 (36)	202 (41)	1163 (48)	.	96 (25)	.
Employed (paid job)	479 (70)	368 (76)	1274 (52)	433 (72)	325	.
Back pain characteristics						
Acute back pain (<3 months)	228 (33)	.	816 (33)	605 (100)	.	174 (50)
History of back pain	549 (80)	.	.	.	358 (94)	.
Severity of back pain (NRS), mean (SD)	6.6 (2.0)	3.9 (2.3)	3.9 (2.6)	5.6 (1.8)	.	5.9 (2.7)
Pain radiating in the leg below the knee	450 (66)	299 (61)	897 (36)	.	163 (45)	137 (39)
Continuous back pain	347 (51)	.	.	.	134 (35)	.
Disability (RDQ), mean (SD)	13.5 (5.2)	8.5 (5.9)	8.8 (6.3)	15.8 (4.7)	13.2 (...)	12.4 (5.8)
History of back surgery	112 (16)	.	.	0 (0)	23 (6)	.

Data are presented as numbers (percentages) unless otherwise indicated
LBP: low back pain; SD: standard deviation; NRS: numeric rating scale (range 0-10, 0 means no pain); RDS: Roland disability questionnaire (range 0-24), a higher score indicates worse health

statement, studies are needed to reveal the reasons why general practitioners refer patients for MRI; such studies will elucidate the decision-making process of the general practitioners. Unfortunately, in our cohort, no data were collected on the specific reasons for referral of patients. Therefore, in the future, we recommended to evaluate and record all reasons for referral, including reasons such as 'reassuring' the patient, or with the aim to avoid unnecessary hospital referrals.

Diagnostic accuracy

In April 2010 the National Health Care Institute of the Netherlands established that health care insurers are obliged to also insure MRI if it is requested by the general practitioner [15]. Currently, most insurance companies reimburse MRI for the lumbar spine in primary care, and there are signs that it is widely used.

On the other hand, a recent update of the guideline of the Dutch College of General Practitioners (the NHG) recommends no advanced imaging in patients with low back pain in general practice at all, but referral to a specialist when imaging is needed [16, 17]. They give several reasons for this recommendation. First, the workgroup that drafted the guideline finds that a significant proportion of general practitioners has difficulty interpreting the MRI findings. It says it is, therefore, not to be expected that MRI will result in a decrease in the number of referrals to secondary care. Second, the workgroup believes it is conceivable that, when MRI diagnostics are available in primary care, MRI is requested more often and earlier in the course of the complaints, possibly at the request of the patient. And it has never been demonstrated that early imaging has a positive influence on the course of the symptoms or reduction of patient anxiety. Third, the workgroup believes that it is likely that the high prevalence of findings that are of no clinical consequence will lead to an increase in experience of illness and anxiety in the patient.

In this way the NHG distances itself from the decision by the National Health Care Institute to give general practitioners the opportunity to refer low back pain patients for MRI themselves. Is this a good development or do we fall short when we withhold MRI from general practitioners?

The ultimate goal of any diagnostic test is to improve the clinical outcome of the patient. From several randomized controlled trials we know that immediate routine lumbar-spine imaging in patients with low back pain and no features suggesting a serious underlying condition, does not improve

clinical outcomes compared with usual clinical care without immediate imaging [18]. These conclusions are mainly based on studies including patients with acute or subacute, non-specific low back pain. Nevertheless, in the present study **(Chapter 2)** we observed that the general practitioners in our cohort mostly referred patients with long-lasting severe back complaints, and/or severe sciatica. To date there are no studies on the diagnostic accuracy of MRI in this specific group of patients. Although we now know that MRI reports do not provide additional prognostic value regarding recovery (**Chapter 3**), we can conclude that the real diagnostic accuracy of lumbar MRI in this group of patients is still unknown.

The use of MRI by general practitioners to diagnose and manage patients with low back pain might indeed lead to more false-positive findings and increased medical costs. But on the other hand, there is a possibility that it could also lead to a faster diagnosis, better reassurance of the patients, and might even avoid unnecessary hospital referrals. In **Chapter 4** we noted that 55% of the patients in our cohort were referred to a specialist after MRI. Thus, there is a possibility that the use of MRI by the general practitioner resulted in less referrals to the hospital in up to 45% of the patients. In order to evaluate whether MRI of the lumbar spine should enter the diagnostic pathway in primary care through early access for general practitioners, or be restricted to secondary care at the request of specialists, a well-designed randomized trial on the diagnostic accuracy and cost-effectiveness is required.

Red flags

One of the other main reasons to refer a patient with low back pain for imaging is when there is suspicion of serious pathology (fracture, malignancy and discitis). Identification of serious pathologies (when they exist) is important in clinical assessment and further evaluation and, usually, specific treatment is required, particularly for malignancy [19, 20]. Despite the potential consequences of a late or missed diagnosis of these serious pathologies, their low prevalence in primary care settings does not justify routine ancillary testing of patients presenting with low back pain [21]. For this reason, accurate screening tools to aid clinical decisions as to when to refer for further testing are essential. Most clinical practice guidelines for back pain recommend the use of red flags to help identify those patients with a higher likelihood of spinal fracture or malignancy, who then become candidates for more extensive diagnostic investigations [17, 21, 22].

However, because various guidelines have produced different lists of red flags to screen for spinal fracture and malignancy, this has led to confusion. None of the guidelines endorse the same single set of red flags for either condition; this means that, for clinicians, it remains unclear what red flags they should use in clinical care [23]. For example, a red flag such as 'back pain started after the age of 50 years' is often seen by general practitioners as a warning signal. In our study **(Chapter 2)**, malignancy was reported in 5 patients, all of whom reported the red flag 'back pain started after the age of 50 years'. However, despite this, of the 8 studies on the diagnostic performance of red flags for malignancy in low back pain [24], none included this particular red flag. Furthermore, no study presented data on a combination of red flags to identify spinal malignancy.
The diagnostic accuracy of red flags for all serious diseases remains a challenging topic for future research on low back pain. Recently, Dutch researchers were able to access associations between red flags and vertebral fractures in primary care [25]. Age ≥75 years, trauma, osteoporosis, a back pain intensity score ≥7, and thoracic pain were associated with a higher chance of getting the diagnosis of a vertebral fracture. In our opinion, an evaluation is also needed of the performance of different combinations of red flags to screen for malignancy, including all relevant red flags currently used in general practice.

Lumbar spinal stenosis
Sciatica, also known as lumbosacral radicular syndrome, is characterized by radiating pain in the leg in the area served by one or more of the spinal nerve roots. About 90% of the time, sciatica is due to a spinal disc herniation pressing on one of the lumbar or sacral nerve roots. Another cause of sciatica is stenosis of the spinal canal. Lumbar spinal stenosis is commonly used to describe patients with symptoms related to an anatomic reduction of the lumbar spinal canal size [26]. This spinal stenosis arises especially with increasing age due to the degenerative spine. One of the radiological tests currently used to diagnose lumbar spinal stenosis is MRI.

In our cohort study, the MRI reports showed spinal stenosis in only 13% of the cases **(Chapter 2)**. This prevalence is much lower than reported in other cohorts including low back pain patients (Table 2). One reason for this could be that the general practitioners in our cohort decided to refer patients likely to have spinal stenosis to a specialist instead of referring them for MRI. Another reason could be the fact that there is no generally

Table 2 Baseline MRI characteristics of patients in low back pain cohorts

	de Schepper et al. [Chapter 2] n=683	Jensen et al. [14] N=98	Modic et al. [15] N=150	Modic et al. [15] N=96	Hancock et al. [17] N=30	You et al. [16] N=647	Jarvik et al. [12] N=380	Jarvik et al. [13] N=349
Study characteristics	Referred by general practitioner for MRI	Without LBP	Acute LBP, primary and secondary care	Acute radiculopathy, primary and secondary care	Acute LBP with centralization, physiotherapist	Referred in primary care for MRI	Referred by physician for radiography	Older adults referred for early MRI/CT
Patient characteristics								
Age in years, mean (SD)	49.9 (13)	42.3 (.)	42.7 (11)	43.7 (10)	36.8 (7.4)	49.5 (15)	53.2 (15)	72.8 (6.0)
Male	365 (53)	50 (51)	61 (41)	43 (45)	16 (53)	320 (50)	169 (44)	120 (34)
MRI findings						**(N=448)**	**(N=178)**	
No findings	44 (6)	.	.	.	.	5 (1.1)	.	.
Bulging	308 (45)	51 (52)	.	.	.	.	106 (60)	.
Disc herniation	492 (72)	27 (28)	85 (57)	62 (65)	23 (75)	246 (55)	58 (33)	.
Nerve root compression	472 (69)	.	40 (27)	44 (46)	.	.	13 (7)	.
Spinal stenosis	87 (13)	.	26 (17)	29 (30)	.	217 (48)	35 (20)	.
Spondylolisthesis	56 (8)	.	.	.	.	.	.	.
Serious pathology*	22 (3)	.	.	.	.	12 (2.7)	9 (5)	14 (4)

Data are presented as numbers (percentages) unless otherwise indicated. LBP: low back pain; SD: standard deviation
* Serious pathology: fracture, malignancy and/or discitis

accepted gold standard for the diagnosis of lumbar spinal stenosis [27, 28]. In our cohort, because we used the MRI reports, we relied on diagnostic radiology. Thus, due to the use of MRI reports instead of standardized scoring of the MR images, the presence of spinal stenosis might be underestimated. However, the challenge to this anatomically based scoring definition is that while necessary for the diagnosis of lumbar spinal stenosis, it is not sufficient to determine the severity of symptoms that leads a patient to seek treatment [26]. The extent of narrowing of the spinal canal correlates poorly with symptom severity, and radiological significant lumbar stenosis can be found in asymptomatic individuals [26, 29-31].

Due to poor correlation between the extent of narrowing and symptom severity, correlating symptoms and physical examination findings with imaging results is necessary to establish a definitive diagnosis [26]. Nowadays, a wide range of clinical, electrodiagnostic, and radiological tests are used to diagnose lumbar spinal stenosis. It is important to know the diagnostic value of these tests because false-positive test results may lead to unnecessary surgery and/or costly or invasive additional diagnostic interventions. In **Chapter 5** we systematically reviewed the diagnostic accuracy of tests for the assessment of lumbar spinal stenosis. We conclude that, at present, MRI is the most promising imaging test for lumbar spinal stenosis, avoiding myelography and electrodiagnostic studies. However, because the extent of narrowing of the spinal canal correlates poorly with symptom severity, it is necessary to correlate the imaging results with symptoms and physical examination findings.

In studies on the diagnostic accuracy of tests in detecting lumbar spinal stenosis, there is considerable heterogeneity in the definition of lumbar spinal stenosis, i.e. the definition of lumbar spinal stenosis was often unclear or not specified at all. Other studies have also shown that there is a need for consensus on the criteria used to define and classify lumbar spinal stenosis [27, 28]. A vague definition of an illness, and imprecise criteria to either rule in or rule out an illness, pose a major problem when performing research in patients with such a disorder [28]. In the absence of widely accepted diagnostic criteria, studies will devise their own construct and this limits the generalizability of the findings. Further research on the use of diagnostics in lumbar spinal stenosis is essential. However, at a time when other musculoskeletal disease experts are considering revisions of well-established sets of criteria [32, 33], the absence of diagnostic and/

or classification criteria in the field of lumbar spinal stenosis should be considered a major focus for international organizations and clinical investigators. Recent studies on clinical tests used the consensus diagnosis of multiple expert spine clinicians as reference standard. However, this raises a problem related to incorporation bias, whereby the overall clinical findings are taken into account in establishing the diagnosis. However, because a diagnosis of the clinical syndrome of lumbar spinal stenosis requires information from the clinical examination, we think that such a bias is unavoidable.

Lumbar disc degeneration

Once patients with radiculopathy and serious causes of back pain (such as fracture and malignancy) are excluded, the remaining patients (about 90%) are those with so-called 'non-specific low back pain'. In these patients, clinicians tend to apply generic symptomatic treatments, such as advice to stay active and avoid bed rest, and the use of analgesic medicines, exercise and manipulation. Although this approach is relatively simple, it does not always work well. The limitations of current approaches are illustrated by many systematic reviews of treatments for low back pain that reveal that existing treatments for non-specific low back pain have, at best, only small effects [34-36]. One reason for this might be that the 'one-size fits all approach' advocated by guidelines fails to target treatments at patients who might benefit the most, thereby 'diluting' their potential benefits [37]. Identifying subgroups of patients for whom different treatments are superior has been referred to as the 'Holy Grail' of low back pain research. One of the possible subgroups based on pathoanatomical findings are patients with severe degeneration of the spine. In the second part of this thesis, we performed an epidemiological study to investigate the characteristics of lumbar disc degeneration.

Lumbar spine osteoarthritis

One of the most common findings on radiography is lumbar disc degeneration, which is characterized by the presence of osteophytes, endplate sclerosis, and disc space narrowing. Large population-based studies have consistently demonstrated that disc degeneration is associated with low back pain [38, 39]. In **Chapter 6**, we contributed to the evidence by reporting different (possible) definitions of lumbar disc degeneration and their association with low back pain. Disc space narrowing at 2 or more

levels appeared to be more strongly associated with low back pain than osteophytes, and that association was only slightly less than the association of pain and radiologic knee osteoarthritis [40] and slightly more than the association between pain and radiological hand osteoarthritis [41] in the same open population sample.
However, the association between low back pain and degenerative changes in the lumbar spine remains complex. In the spine, the presence of both disc degeneration and osteophyte formation at the same vertebral level has been used to define lumbar spine osteoarthritis [42]. Estimates of the presence of lumbar spine osteoarthritis are high, ranging from 40-85%; however, this large range is primarily due to differences in definitions [43, 44]. The definition is debatable because it lacks the anatomical synovial structures necessary to meet the definition of osteoarthritis. Therefore, one of the other definitions used to define lumbar spine osteoarthritis includes the facet joint, which is a synovial joint.

At every spinal level, the paired facet joint and the intervertebral disc make up the 'three-joint complex', or the spinal 'motion segment'. The facet joint itself tends to degenerate in concert with the disc [45]. Suri et al. [46] observed degeneration of the spine with an ordered sequence beginning in the anterior structures (disc degeneration) which, in some cases, was followed by degeneration of the posterior joint (facet joints). However, 22% of individuals demonstrated an atypical pattern of degeneration, beginning in the posterior joints. Increased age and body mass index, as well as female gender, were related to this facet joint osteoarthritis. Other studies have been unable to link facet osteoarthritis with low back pain [47-49], except for the study of Suri et al. [50] which shows that the presence and extent of severe facet joint osteoarthritis was associated with back pain in community-based older adults, independent of sociodemographics, health factors, and disc height narrowing. Furthermore, there are indications that facet joint degeneration may be associated with osteoarthritis in other joints [51]. Goode et al. [52] found a significant association between radiographic knee osteoarthritis and hand osteoarthritis and facet joint osteoarthritis; they found no significant association between disc degeneration and hip, knee or hand osteoarthritis, or between vertebral osteophytes and hip and hand osteoarthritis.

Findings such as these suggest that facet joint osteoarthritis may have a role in refining back pain case definition or directing back pain treatment

for older adults. However, a complicating factor is that facet osteoarthritis is best assessed with CT which, unfortunately, contains too much radiation to use in research or clinical practice. As such, a clinically relevant definition that combines spine features to accurately represent spine osteoarthritis is needed.

Osteoarthritis of the knee and hip has clinical classification criteria, as described by the American College of Rheumatology (ACR). The ACR criteria state that, besides pain, morning stiffness is an important criteria for hip and knee osteoarthritis [53, 54]. Therefore, in **Chapter 7** we explored the association between spinal morning stiffness and lumbar disc degeneration. We found a moderate association between the two, and the magnitude of the association was similar to the association between morning stiffness in the legs and knee or hip osteoarthritis. Unfortunately, because of the use of lateral radiographs, we were unable to examine if the presence of facet joint osteoarthritis was responsible for this association, or whether lumbar disc degeneration was associated with morning stiffness independently of facet joint osteoarthritis. Our results might indicate that spinal morning stiffness is one of the symptoms that clinicians could use for sub-grouping low back pain patients with symptoms due to lumbar spine osteoarthritis.

Spine disorders versus hip disorders

In **Chapter 8** we studied the possible association between osteoarthritis of the spine and hip pain. Differentiating back pain from hip pain in patients who present with classic signs and symptoms is mostly not difficult. However, in some cases, patients present with nonspecific complaints of pain in the buttock, lateral hip or thigh. The differential diagnosis of this nonspecific pain is broad and includes radiating pain from the lumbar spine.

The differentiation of signs and symptoms suggestive of hip disorders versus spine disorders is important in giving patients the most beneficial treatment, especially if the treatment includes a major reconstructive surgery, such as hip replacement. In order to make a differentiation between hip and spine originated pain there have been a few studies about the usefulness of local anesthetic hip infiltrations [55]. However, the available evidence had a high risk of bias, and no recommendation could be made regarding substantiated favoring or not favoring the use of intra-articular injections for the diagnosis of hip osteoarthritis.

To our knowledge, there have been no studies about the usefulness of local spine infiltrations to differentiate hip and spine originated hip pain. However, infiltration of every patient with atypical hip pain for possible coexistent lumbar spine osteoarthritis would be counterproductive and costly. Preoperative identification of factors associated with hip pain arising from the lumbar spine would aid the physician by identifying the subgroup of patients who might not experience full relief of pain with a hip arthroplasty.
Our data in **Chapter 8** provided evidence for radiating pain from the degenerated higher lumbar spine (L1/L2 and L2/L3) as a possible cause of hip pain. We think that in case of uncertainty of the cause of hip pain, evaluation of lumbar radiographs may help to identify those hip pain patients who might have pain arising from the lumbar spine and, therefore, will be unlikely to benefit from hip replacement. However, additional well-designed studies are needed to verify this hypothesis.

Implications for clinical practice
The question remains: is the subgroup of patients with lumbar spine osteoarthritis a group for whom different treatments are superior? Currently, our understanding of the conservative treatment for low back pain is, stated simply, that some activity is better than no activity [42, 56, 57]. Exercise therapy is an activity which has long been a treatment option for low back pain, with Cochrane and other review indicating some effectiveness for treating chronic low back pain [58]. The use of conservative treatment to prevent spine osteoarthritis or conservative treatment for spine osteoarthritis as a primary technique for treating low back pain has not yet been reported in the literature [42].
Furthermore, there are secondary intervention techniques specifically aimed at treating symptomatic spine osteoarthritis. Facet joint injection therapy to treat low back pain has increased dramatically in recent years. In the United States, intervention for facet joint pain has increased substantially with annual growth of 60% from 1997 to 2006 [59]. However, there is little evidence to support the use of this therapy, with recent practice guidelines recommending against the use of facet joint steroid injections [42, 60].
Furthermore, joint replacement surgery is currently not available for severe lumbar osteoarthritis, unlike knee or hip replacement that is available for severe knee or hip osteoarthritis.

Therefore, also in patients with lumbar osteoarthritis, generic symptomatic treatments such as advice to stay active, analgesic medicines and patient education are recommended. And as a result, the use of lumbar radiography to diagnose lumbar osteoarthritis is at present not recommended in clinical practice.

Implications for research

Most research has focused on treating symptoms and functional impairment of low back pain rather than on understanding the mechanisms underlying the anatomic and functional changes we currently call spine osteoarthritis. But because there is no evidence that lumbar imaging improves outcomes for patients with low back pain this does not necessarily mean that we should not aim for a better understanding of the pathology identified on imaging. To dismiss investigations on imaging (which aim to elucidate the source and causes of low back pain) because imaging is currently not recommended for low back pain is to miss the point of this line of research [34].

We think that besides identifying subgroups based on predicted chronicity or pain phenotypes, we should also focus on identifying subgroups based on pathoanatomical findings. In other words, there is more need for research that aims to identify methods and tests that allow clinicians to determine the origin of a patient's back pain. A good example of this is the study of van Hoeven et al. [61, 62]; the authors aimed to design a simple referral model for general practitioners that would identify patients at risk for axial spondyloarthritis.

A common argument against this is that, currently, there is no evidence that diagnosis improves patients' outcomes [19]. A possible reason for this is the fact that it is not possible for a diagnosis to influence patients' outcomes if no effective treatment exists for the specific disease or pathology identified. A diagnosis may be of value even without the availability of effective treatment, as it may provide a logical pathway for the development and testing of future interventions [34]. A better understanding of the pathological source of low back pain is likely to precede the identification of new and effective treatments for low back pain.

References

1. Andersson HI, Ejlertsson G, Leden I, Schersten B. Impact of chronic pain on health care seeking, self care, and medication. Results from a population-based Swedish study. *J Epidemiol Community Health*. 1999 Aug;53(8):503-9.
2. Molano SM, Burdorf A, Elders LA. Factors associated with medical care-seeking due to low-back pain in scaffolders. *Am J Ind Med*. 2001 Sep;40(3): 275-81.
3. Waxman R, Tennant A, Helliwell P. Community survey of factors associated with consultation for low back pain. *BMJ*. 1998 Dec 5;317(7172):1564-7.
4. Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am*. 2012 Jul;50(4):569-85.
5. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ*. 2007 Jun 23;334(7607):1313-7.
6. Campbell P, Foster NE, Thomas E, Dunn KM. Prognostic indicators of low back pain in primary care: five-year prospective study. *J Pain*. 2013 Aug;14(8): 873-83.
7. Hill JC, Konstantinou K, Egbewale BE, et al. Clinical outcomes among low back pain consulters with referred leg pain in primary care. *Spine (Phila Pa 1976)*. 2011 Dec 1;36(25):2168-75.
8. Mehling WE, Gopisetty V, Bartmess E, et al. The prognosis of acute low back pain in primary care in the United States: a 2-year prospective cohort study. *Spine (Phila Pa 1976)*. 2012 Apr 15;37(8):678-84.
9. Jarvik JG, Hollingworth W, Martin B, et al. Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *JAMA*. 2003 Jun 4;289(21):2810-8.
10. Jarvik JG, Gold LS, Comstock BA, et al. Association of Early Imaging for Back Pain With Clinical Outcomes in Older Adults (vol 313, pg 1143, 2015). *Jama-J Am Med Assoc*. 2015 May 5;313(17):1758-.
11. Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med*. 1994 Jul 14;331(2):69-73.
12. Modic MT, Obuchowski NA, Ross JS, et al. Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. *Radiology*. 2005 Nov;237(2):597-604.
13. You JJ, Bederman SS, Symons S, et al. Patterns of care after magnetic resonance imaging of the spine in primary care. *Spine (Phila Pa 1976)*. 2013 Jan 1; 38(1):51-9.

14. Hancock M, Maher C, Macaskill P, et al. MRI findings are more common in selected patients with acute low back pain than controls? *Eur Spine J*. 2012 Feb;21(2):240-6.
15. (CVZ) CvZ. Standpunt Aanvraag MRI door een huisarts in de Zorgverzekering-swet. Diemen2010.
16. Annemieke Schaafstra WS, Sonja Bons, Maarten Borg, Bart Koes, Raymond Ostelo, Antje Spijker-Huiges, Jako Burgers, Margriet Bouma, Arianne Verburg. NHG-Standaard Lumbosacraal radiculair syndroom (Tweede herziening) *Huisarts & Wetenschap*. 2015;58(6):308-20.
17. Chavannes AW MJ, Koes BW, et al. NHG-Standaard Aspecifieke lagerugpijn. *Huisarts Wet*. 2005;48:113-23.
18. Chou R, Fu R, Carrino JA, Deyo RA. Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet*. 2009 Feb 7;373(9662):463-72.
19. Chou R, Qaseem A, Snow V, et al. 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*. 2007 Oct 2;147(7):478-91.
20. Henschke N, Maher CG, Refshauge KM, et al. Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. *Arthritis Rheum*. 2009 Oct;60(10):3072-80.
21. Downie A, Williams CM, Henschke N, et al. Red flags to screen for malignancy and fracture in patients with low back pain: systematic review. *BMJ*. 2013;347: f7095.
22. Mens JMA CA, Koes BW, Lubbers WJ, Ostelo RWJG, Spinnewijn WEM, Kolnaar BGM. . NHG-Standaard Lumbosacraal radiculair syndroom (LRS). *Huisarts Wet*. 2005;48(4):171-8.
23. Koes BW, van Tulder M, Lin CW, et al. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J*. 2010 Dec;19(12):2075-94.
24. Henschke N, Maher CG, Ostelo RW, et al. Red flags to screen for malignancy in patients with low-back pain. *Cochrane Database Syst Rev*. 2013;2:CD008686.
25. Enthoven WT, Geuze J, Scheele J, et al. Prevalence and Red Flags Regarding Specified Causes of Back Pain in Older Adults Presenting in General Practice. *Phys Ther*. 2015 Jul 16.
26. Genevay S, Atlas SJ. Lumbar spinal stenosis. *Best Pract Res Clin Rheumatol*. 2010 Apr;24(2):253-65.
27. Genevay S, Atlas SJ, Katz JN. Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: a structured literature review. *Spine (Phila Pa 1976)*. 2010 Apr 1;35(7):803-11.

28. Steurer J, Roner S, Gnannt R, et al. Quantitative radiologic criteria for the diagnosis of lumbar spinal stenosis: a systematic literature review. *BMC Musculoskelet Disord*. 2011;12:175.
29. Boden SD, Davis DO, Dina TS, et al. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am*. 1990 Mar;72(3):403-8.
30. Kalichman L, Cole R, Kim DH, et al. Spinal stenosis prevalence and association with symptoms: the Framingham Study. *Spine J*. 2009 Jul;9(7):545-50.
31. Wiesel SW, Tsourmas N, Feffer HL, et al. A study of computer-assisted tomography. I. The incidence of positive CAT scans in an asymptomatic group of patients. *Spine (Phila Pa 1976)*. 1984 Sep;9(6):549-51.
32. Liao KP, Batra KL, Chibnik L, et al. Anti-cyclic citrullinated peptide revised criteria for the classification of rheumatoid arthritis. *Ann Rheum Dis*. 2008 Nov;67(11):1557-61.
33. Taylor W, Gladman D, Helliwell P, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. *Arthritis Rheum*. 2006 Aug;54(8):2665-73.
34. Hancock MJ, Maher CG, Laslett M, et al. Discussion paper: what happened to the 'bio' in the bio-psycho-social model of low back pain? *Eur Spine J*. 2011 Dec;20(12):2105-10.
35. Deyo RA. Treatments for back pain: can we get past trivial effects? *Ann Intern Med*. 2004 Dec 21;141(12):957-8.
36. Machado LA, Kamper SJ, Herbert RD, et al. Analgesic effects of treatments for non-specific low back pain: a meta-analysis of placebo-controlled randomized trials. *Rheumatology (Oxford)*. 2009 May;48(5):520-7.
37. Foster NE, Hill JC, Hay EM. Subgrouping patients with low back pain in primary care: are we getting any better at it? *Man Ther*. 2011 Feb;16(1):3-8.
38. Raastad J, Reiman M, Coeytaux R, et al. The association between lumbar spine radiographic features and low back pain: a systematic review and meta-analysis. *Semin Arthritis Rheum*. 2015 Apr;44(5):571-85.
39. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine (Phila Pa 1976)*. 1997 Feb 15;22(4):427-34.
40. Claessens AA, Schouten JS, van den Ouweland FA, Valkenburg HA. Do clinical findings associate with radiographic osteoarthritis of the knee? *Ann Rheum Dis*. 1990 Oct;49(10):771-4.
41. Dahaghin S, Bierma-Zeinstra SM, Ginai AZ, et al. Prevalence and pattern of radiographic hand osteoarthritis and association with pain and disability (the Rotterdam study). *Ann Rheum Dis*. 2005 May;64(5):682-7.

42. Goode AP, Carey TS, Jordan JM. Low back pain and lumbar spine osteoarthritis: how are they related? *Curr Rheumatol Rep*. 2013 Feb;15(2):305.
43. A.M.J. Chorus DFS. Nationale Peiling Bewegingsapparaat 2010. TNO-rapport2011.
44. Battie MC, Videman T, Parent E. Lumbar disc degeneration: epidemiology and genetic influences. *Spine (Phila Pa 1976)*. 2004 Dec 1;29(23):2679-90.
45. Gellhorn AC, Katz JN, Suri P. Osteoarthritis of the spine: the facet joints. *Nat Rev Rheumatol*. 2013 Apr;9(4):216-24.
46. Suri P, Miyakoshi A, Hunter DJ, et al. Does lumbar spinal degeneration begin with the anterior structures? A study of the observed epidemiology in a community-based population. *BMC Musculoskelet Disord*. 2011;12:202.
47. Kalichman L, Li L, Kim DH, et al. Facet joint osteoarthritis and low back pain in the community-based population. *Spine (Phila Pa 1976)*. 2008 Nov 1; 33(23):2560-5.
48. Kjaer P, Leboeuf-Yde C, Korsholm L, et al. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine (Phila Pa 1976)*. 2005 May 15;30(10):1173-80.
49. Savage RA, Whitehouse GH, Roberts N. The relationship between the magnetic resonance imaging appearance of the lumbar spine and low back pain, age and occupation in males. *Eur Spine J*. 1997;6(2):106-14.
50. Suri P, Hunter DJ, Rainville J, et al. Presence and extent of severe facet joint osteoarthritis are associated with back pain in older adults. *Osteoarthritis Cartilage*. 2013 Sep;21(9):1199-206.
51. Borenstein D. Does osteoarthritis of the lumbar spine cause chronic low back pain? *Curr Rheumatol Rep*. 2004 Feb;6(1):14-9.
52. Goode AP, Marshall SW, Renner JB, et al. Lumbar spine radiographic features and demographic, clinical, and radiographic knee, hip, and hand osteoarthritis. *Arthritis Care Res (Hoboken)*. 2012 Oct;64(10):1536-44.
53. Altman R, Alarcon G, Appelrouth D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum*. 1991 May;34(5):505-14.
54. Altman R, Asch E, Bloch D, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum*. 1986 Aug;29(8):1039-49.
55. Dorleijn DM, Luijsterburg PA, Bierma-Zeinstra SM, Bos PK. Is anesthetic hip joint injection useful in diagnosing hip osteoarthritis? A meta-analysis of case series. *J Arthroplasty*. 2014 Jun;29(6):1236-42 e1.

56. Bigos SJ, Holland J, Holland C, et al. High-quality controlled trials on preventing episodes of back problems: systematic literature review in working-age adults. *Spine J*. 2009 Feb;9(2):147-68.
57. Buchbinder R, Pransky G, Hayden J. Recent advances in the evaluation and management of nonspecific low back pain and related disorders. *Best Pract Res Clin Rheumatol*. 2010 Apr;24(2):147-53.
58. Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database Syst Rev*. 2005(3):CD000335.
59. Manchikanti L, Singh V, Pampati V, et al. Analysis of growth of interventional techniques in managing chronic pain in the Medicare population: a 10-year evaluation from 1997 to 2006. *Pain Physician*. 2009 Jan-Feb;12(1):9-34.
60. Chou R, Atlas SJ, Stanos SP, Rosenquist RW. Nonsurgical interventional therapies for low back pain: a review of the evidence for an American Pain Society clinical practice guideline. *Spine (Phila Pa 1976)*. 2009 May 1;34(10): 1078-93.
61. van Hoeven L, Luime J, Han H, et al. Identifying axial spondyloarthritis in Dutch primary care patients, ages 20-45 years, with chronic low back pain. *Arthritis Care Res (Hoboken)*. 2014 Mar;66(3):446-53.
62. van Hoeven L, Vergouwe Y, de Buck PD, et al. External Validation of a Referral Rule for Axial Spondyloarthritis in Primary Care Patients with Chronic Low Back Pain. *PLoS One*. 2015;10(7):e0131963.

Summary

Low back pain is a major health problem. A recent publication showed that from the 289 diseases and conditions that were investigated, low back pain leads to the most 'years lived with disability'. It is one of the most frequently occurring musculoskeletal disorders and results in high costs due to disability, lost time from work, and medical treatment and care.

Low back pain is usually divided in to 'specific' and 'non-specific' low back pain. Specific low back pain is defined as symptoms caused by a particular pathophysiological mechanism, such as hernia nuclei pulposi, discitis, malignancy or fracture. Non-specific low back pain is defined as pain for which no specific cause can be shown. At about 80 to 95% of people with low back pain, no specific cause is found. Overall, adequate treatment of patients with low back pain begins with making the correct diagnosis. In the Netherlands, the first step in diagnostics patients with low back pain is mostly made by the general practitioner, almost exclusively based on history and physical examination. In addition, in a small proportion of patients, diagnostic imaging is needed. International guidelines recommend the use of imaging only when there is suspicion of serious pathology (fracture, malignancy and discitis), or in patients with severe sciatica for whom surgery is indicated because they fail to respond to conservative care for at least 6 to 8 weeks. Currently, there is widespread consensus that there is no indication to perform diagnostic imaging in patients with non-specific low back pain.

MRI in primary care

If a physician suspects the presence of a specific disease, diagnostic imaging can be used. In recent years, general practitioners in the Netherlands can refer low back pain patients for MRI of the lumbar spine themselves. Possible reasons for the use of MRI in general practice are i) to detect or exclude specific pathologies, ii) to reassure the patient (and physician), iii) and/or to prevent unnecessary referrals to secondary care. Despite the recommendations of the guidelines to use MRI only in specific cases, the use of MRI as the initial imaging for low back pain seems to become more common in general practice.

However, data on the use of MRI by general practitioners in the Netherlands was still lacking. Therefore, we designed an observational prospective cohort study with a 12-month follow-up in low back pain patients referred

for MRI in general practice. Patients referred by their general practitioner for MRI of the lumbar spine were recruited at the MRI Center. The MRI radiology reports were scored and information on patients' characteristics, characteristics of the complaint, and red flags were derived at baseline and at 3 and 12-months follow-up.

In **Chapter 2**, we explored the characteristics and MRI findings of the patients with low back pain referred for MRI in general practice. A total of 683 patients (53% male) were included, with a mean age of 49.9 (range 19-80 years). The mean back pain severity was 6.6 (SD 2.0) and 67% of the patients reported having chronic low back pain. In total, 69% of the MRI reports mentioned signs of nerve root compression. Serious pathologies (fracture, malignancy, discitis) were reported in 3% of the patients. Patients with malignancy were older and reported less often a history of back pain complaints. These results showed us that a large proportion of the referred patients in our cohort had complaints for which imaging could be indicated. Specifically, 374 (55%) reported sciatica complaints for at least 6 weeks and could therefore be candidates for surgery in which case imaging is recommended. These results are an indication that general practitioners in the Netherlands tend to adhere to their guidelines. However, to validate this statement, studies are needed to reveal the reasons why general practitioners refer patients for MRI.

The ultimate goal of any diagnostic test is to improve the clinical outcome of the patient. Currently, well-designed randomized trials provide the most accurate diagnostic evidence, because they provide the most direct information about the clinical benefits and harms of alternative testing strategies. However, in daily practice most studies on various diagnostic tests estimate how accurately they can identify a disease or condition, or how well the test provides prognostic information. Furthermore, identification of prognostic factors predicting recovery, persistent pain, and disability are important for better understanding of the clinical course, to inform patients and physicians and support therapeutic decision making.

In **Chapter 3**, we investigated the added prognostic value of baseline MRI findings over known prognostic factors for recovery. Multivariate logistic regression analysis was performed in 3 steps: derivation of a predictive model including characteristics of the patients and back pain only (history taking), including reported MRI findings only, and the addition of

reported MRI findings to the characteristics of the patients and back pain. At 12-months follow-up 53% of the patients reported recovery. Lower age, better attitude/beliefs regarding back pain, acute back pain, presence of neurological symptoms of the leg(s), and presence of non-continuous back pain were significantly associated with recovery at 12-months follow-up: area under the curve (AUC) 0.77. Addition of the MRI findings did not resulted in a stronger prediction of recovery. In a clinical perspective, these results raise some questions about the usefulness of lumbar MRI in general practice. But in the end only a well-conducted randomized trial on the diagnostic accuracy and cost-effectiveness can give us more information.

The use of MRI by general practitioners to diagnose and manage patients with low back pain could possibly result in avoidance of unnecessary hospital referrals. However, data on the patterns of subsequent care among patients referred for lumbar spine MRI by their GP are scarce. A better understanding of the patterns of healthcare services used after MRI of the lumbar spine would provide information about how MRI findings are used by GPs for subsequent management. Identification of possible prognostic factors predicting consultation with specialists or surgery can be important to inform patient and physician. Therefore, in **Chapter 4**, we investigated the association between patient characteristics, back pain characteristics and MRI abnormalities with subsequent specialist consultation and/or surgery in the same low back pain patients referred to MRI by their general practitioner. Of the 683 included patients, 301 (55%) reported consultation with a specialist during the first 3 months, and 124 (18%) underwent spine surgery during 12 months of follow-up. Younger patients, with pain radiating in the leg below the knee, severe disability, a history of back surgery, presence of nerve root compression or spinal stenosis on MRI were more likely to undergo subsequent spine surgery (AUC 0.75).

Lumbar spinal stenosis

One of the anatomical abnormalities that can be found with MR imaging is lumbar spinal stenosis. Lumbar spinal stenosis is commonly used to describe patients with symptoms related to an anatomic reduction of the lumbar spinal canal size. The extent of narrowing of the spinal canal correlates poorly with symptom severity and radiological significant lumbar stenosis can be found in asymptomatic individuals. As a consequence, correlating symptoms and physical examination findings with imaging

results is necessary to establish a definitive diagnosis. A wide range of clinical, electrodiagnostic, and radiological tests are currently used to diagnose lumbar spinal stenosis. It is important to know the diagnostic value of these tests because false-positive test results may lead to unnecessary surgery and/or expensive or invasive additional diagnostic interventions.

In **Chapter 5**, we updated our previous systematic review on the diagnostic accuracy of tests used to diagnose lumbar spinal stenosis. Our previous systematic review included studies up to March 2004; this review is current up to March 2011. Twenty-two additional articles in addition to the 24 included in the previous review met the inclusion criteria. Combined, this resulted in 20 articles concerning imaging tests, 11 articles evaluating electrodiagnostic tests, and 15 articles evaluating clinical tests.

At present, the most promising imaging test for lumbar spinal stenosis is MRI, avoiding myelography because of its invasiveness and lack of superior accuracy. Electrodiagnostic studies showed no superior accuracy for conventional electrodiagnostic testing compared with MRI. These tests should be considered in the context of those presenting symptoms with the highest diagnostic value, including radiating leg pain that is exacerbated while standing up, the absence of pain when seated, the improvement of symptoms when bending forward, and a wide-based gait.

Lumbar disc degeneration

Once patients with radiculopathy and serious causes of back pain (such as fracture and malignancy) are excluded, the remaining patients (about 85%) are those with so called 'non-specific low back pain'. In these patients, clinicians tend to apply generic symptomatic treatments, such as advice to stay active and avoid bed-rest, and the use of analgesic medicines, exercise and manipulation. Although this approach is relatively simple, it does not always work well. The limitations of current approaches are illustrated by many systematic reviews of treatments for low back pain that reveal that existing treatments for non-specific low back pain have, at best, only small effects.

One reason for this might be that the 'one-size fits all approach' advocated by guidelines fails to target treatments at patients who might benefit the most, thereby diluting their potential benefits. Identifying subgroups of

patients for whom different treatments are superior has been referred to as the 'Holy Grail' of low back pain research. One of the possible subgroups based on pathoanatomical findings are patients with severe degeneration of the spine. In the second part of this thesis, we performed an epidemiological study to investigate the characteristics of lumbar disc degeneration.

Lumbar disc degeneration is characterized (radiographically) by the presence of osteophytes, endplate sclerosis, and disc space narrowing. The association between low back pain and degenerative changes in the lumbar spine is complex. It is known that lumbar disc degeneration can be a possible risk factor for back pain in adults, with odds ratios ranging from 1.3 to 3.2. Nevertheless, due to differences in the definitions used for lumbar disc degeneration, studies on the relation between low back pain and lumbar disc degeneration are difficult to compare. This is why it is still difficult to provide a clinically relevant definition for lumbar disc degeneration.

In **Chapter 6**, we therefore explored the association of the different individual radiographic features of lumbar disc degeneration with self-reported low back pain. In an open population based study (Rotterdam Study), 2819 lumbar radiographs were evaluated for the presence and severity of anterior osteophytes and disc space narrowing. Logistic regression was used to determine the association between different definitions of lumbar disc degeneration for low back pain. Osteophytes were the most frequent radiographic feature observed, with men having the greatest frequency. Disc space narrowing was more frequent in women. Both radiographic features increased in frequency with age. Disc space narrowing appeared more strongly associated with low back pain than osteophytes, especially in men (OR=1.9; 95% CI: 1.4-2.8). Disc space narrowing at 2 or more levels appeared more strongly associated with low back pain than disc space narrowing at only 1 level (OR=2.4; 95% CI: 1.6-3.4). After excluding level L5-S1, the strength of almost all associations increased. These associations are only slightly less than the association of pain and radiologic knee osteoarthritis and even slightly more than the association of pain and radiologic hand osteoarthritis in the same population sample. From our data, a useful case definition for lumbar disc degeneration was deduced; specifically disc space narrowing at 2 or more levels from L1/2 to L4/5. As a result it might be a promising clinically relevant phenotype in genetic and epidemiologic lumbar disc degeneration research.

Low back pain patients with symptoms due to lumbar disc degeneration could be a subgroup within the population of non-specific low back pain patients. Clinical symptoms associated with radiographic lumbar disc degeneration may help identify such patients. Although lumbar disc degeneration cannot be defined as real osteoarthritis because the facet joints are the only synovial joints in the spine, it is often used as a proxy for osteoarthritis of the spine. Clinical osteoarthritis of the knee and hip includes (besides pain) also morning stiffness (ACR criteria).

In **Chapter 7**, we therefore explored the association between spinal morning stiffness and lumbar disc degeneration. These associations are also compared with the associations between morning stiffness in the legs, and knee or hip osteoarthritis. Data from the study in Chapter 6 was used to explore the association between spinal morning stiffness and two definitions of lumbar disc degeneration (i.e., 'osteophytes' and 'narrowing'). Spinal morning stiffness combined with low back pain and its association with lumbar disc degeneration was also analyzed. Similar analyses were performed for knee and hip pain, morning stiffness in the legs, and radiographic knee and hip osteoarthritis in order to compare these associations with those of lumbar disc degeneration. Both definitions of lumbar disc degeneration were associated with spinal morning stiffness (OR=1.3; 95% CI: 1.1-1.6 for 'osteophytes' and OR=1.8; 95% CI: 1.4-2.2 for 'narrowing'). Both the odds ratios increased when spinal morning stiffness was combined with low back pain (OR=1.5; 95% CI: 1.1-2.0 for 'osteophytes' and OR=2.5; 95% CI: 1.9-3.4 for 'narrowing'). The magnitude of the association for the definition 'narrowing' is similar to the association between morning stiffness in the legs and knee osteoarthritis (OR=3.0; 95% CI: 2.1-4.1). This might indicate that spinal morning stiffness is one of the symptoms that clinicians could use for sub-grouping low back pain symptoms due to lumbar disc degeneration.

In **Chapter 8** we studied the possible association between osteoarthritis of the spine and hip pain. Differentiating back pain from hip pain in patients who present with classic signs and symptoms is mostly not difficult. However, in some cases, patients present with nonspecific complaints of pain in the buttock, lateral hip or thigh. The differential diagnosis of this nonspecific pain is broad and includes radiating pain from the lumbar spine. The differentiation of signs and symptoms suggestive of hip disorders versus spine disorders is important in giving patients the most beneficial treat-

ment, especially if the treatment includes a major reconstructive surgery, such as hip replacement. Preoperative identification of factors associated with hip pain arising from the lumbar spine would aid the physician by identifying the subgroup of patients who might not experience full relief of pain with a hip arthroplasty.

In **Chapter 8**, we therefore explored the association of self-reported hip pain with the different individual radiographic features of lumbar disc degeneration in the 2819 scored lumbar radiographs from the Rotterdam study (Chapter 6). The presence of disc space narrowing at level L1/L2 was significantly associated with hip pain in the last month (men OR=2.0; 95% CI: 1.1-3.8 and women OR=1.7; 95% CI: 1.1-2.5). The presence of disc space narrowing at level L2/L3 was only significantly associated with hip pain in women. The strength of the associations increased for self-reported chronic hip pain, especially in men (L1/L2 OR=2.5; 95% CI: 1.3-5.0). The presence of disc space narrowing at the lower levels (L3/L4/L5/S1) was not significantly associated with hip pain. We concluded that in case of uncertainty of the cause of hip pain, evaluation of lumbar radiographs may help to identify those hip pain patients who might have pain arising from the lumbar spine. However, well-designed studies are needed to verify this hypothesis.

Finally, in **Chapter 9**, we reflected on the main findings in this thesis, and elaborated on their implications for clinical practice and research.

Nederlandse samenvatting

Lage rugpijn is een belangrijk gezondheidsprobleem. Uit een recente publicatie blijkt dat van de 289 ziekten en aandoeningen die zijn onderzocht, lage rugpijn tot de meeste 'jaren geleefd met een beperking' leidt. Het is een van de meest voorkomende aandoeningen van het bewegingsapparaat en leidt tot hoge kosten als gevolg van beperkingen, arbeidsongeschiktheid, en medische zorg.
Lage rugpijn is meestal verdeeld in 'specifieke' en 'aspecifieke' lage rugpijn. Specifieke lage rugpijn wordt gedefinieerd als rugpijn veroorzaakt door een bepaald pathofysiologisch mechanisme, zoals een discus hernia, discitis, maligniteit of fractuur. Aspecifieke lage rugpijn wordt gedefinieerd als pijn waarvoor geen specifieke oorzaak kan worden aangetoond. Bij ongeveer 80 tot 95% van de mensen met lage rugpijn wordt geen specifieke oorzaak gevonden.
Adequate behandeling van patiënten met lage rugpijn begint vaak met het maken van de juiste diagnose. In Nederland wordt de eerste stap in de diagnostiek bij patiënten met lage rugpijn meestal gemaakt door de huisarts, bijna uitsluitend op basis van anamnese en lichamelijk onderzoek. Maar in een klein aantal patiënten blijkt uiteindelijk diagnostische beeldvorming nodig te zijn. Internationale richtlijnen raden het gebruik van beeldvormende diagnostiek alleen aan wanneer er een vermoeden is van ernstige pathologie (fractuur, maligniteit en discitis), of bij patiënten met ernstige ischias en een indicatie voor chirurgie omdat ze niet goed reageren op conservatieve zorg voor ten minste 6 tot 8 weken. Momenteel is er brede consensus dat er geen indicatie is voor diagnostische beeldvorming bij patiënten met aspecifieke lage rugpijn.
Ondanks al deze aanbevelingen, wordt beeldvorming van de lumbale wervelkolom (röntgenfoto, CT en MRI) nog vaak uitgevoerd bij patiënten met lage rugpijn en wordt deze ook vaak uitgevoerd in afwezigheid van een duidelijke indicatie. Het doel van het in dit proefschrift beschreven onderzoek is om meer inzicht te verschaffen in de diagnostische instrumenten die gebruikt worden bij patiënten met lage rugpijn. Ten eerste hebben we onderzocht hoe lumbale MRI momenteel wordt gebruikt door huisartsen in Nederland.

MRI in de eerste lijn

Als een arts de aanwezigheid vermoedt van een specifieke ziekte, dan kan diagnostische beeldvorming worden gebruikt. Sinds een aantal jaren kunnen Nederlandse huisartsen patiënten met lage rugpijn verwijzen voor

een MRI van de lumbale wervelkolom. Mogelijke redenen voor het gebruik van MRI in de eerste lijn zijn i) het detecteren of uitsluiten van specifieke pathologie, ii) het geruststellen van de patiënt (en arts), iii) en/of het voorkomen van onnodige verwijzingen naar de tweede lijn. Ondanks de aanbevelingen om MRI alleen te gebruiken in specifieke gevallen, lijkt het gebruik van MRI als de initiële beeldvorming voor lage rugpijn vaker voor te komen in de huisartsenpraktijk. Echter, gegevens over het gebruik van MRI door huisartsen in Nederland ontbraken nog. Dit was voor ons een reden om een observationele, prospectieve cohort studie op te zetten met lage rugpijn patiënten die zijn verwezen voor MRI door hun eigen huisarts. De patiënten werden geïncludeerd bij het MRI Centrum, en werden vervolgens 12 maanden gevolgd. De MRI-rapporten werden gescoord, en informatie over de kenmerken van de patiënt, kenmerken van de klacht, en eventuele rode vlaggen werden verzameld op baseline, en na 3 en 12 maanden follow-up.

In **hoofdstuk 2** deden we onderzoek naar de kenmerken van de patiënten en naar de gevonden MRI bevindingen. In totaal werden 683 patiënten geïncludeerd (53% man), met een gemiddelde leeftijd van 49,9 (spreiding 19-80 jaar). De gemiddelde ernst van de rugpijn was 6.6 (SD 2.0) op een schaal van 0 tot 10, en 67% van de patiënten meldde chronische klachten. In totaal werd er in 69% van de MRI-rapporten melding gemaakt van zenuwwortel compressie. Bij 3% van de patiënten werd er melding gemaakt van een ernstige afwijking (wervelfractuur, maligniteit, discitis). Patiënten met een maligniteit waren significant ouder en rapporteerden minder vaak een geschiedenis van rugklachten. Deze resultaten toonden ons dat een groot deel van de patiënten in ons cohort klachten had waarbij beeldvormende diagnostiek kan worden aangeraden. Specifieker, 374 patiënten (55%) rapporteerden ischias klachten van tenminste 6 weken en hadden dus klachten waarbij beeldvorming kan worden toegepast. Deze resultaten zijn een indicatie dat de huisartsen in Nederland zich houden aan de richtlijn. Echter, om dit te valideren zijn er studies nodig naar de verschillende redenen van huisartsen om patiënten te verwijzen voor MRI.

Het uiteindelijke doel van een diagnostische test is om de klinische uitkomst van de patiënt te verbeteren. Momenteel staan goed uitgevoerde gerandomiseerde studies bovenaan de onderzoeks hiërarchie, maar in de praktijk kijken de meeste studies naar diagnostische tests naar hoe nauwkeurig zij een ziekte of aandoening aantonen, of hoe goed de test prog-

nostische informatie geeft. Bovendien, de identificatie van prognostische factoren welke herstel, aanhoudende pijn en beperkingen voorspellen zijn van belang voor een beter begrip van het klinisch beloop, en om patiënten en artsen te informeren en te ondersteunen bij hun therapeutische besluitvorming.

In **hoofdstuk 3** onderzochten we daarom de toegevoegde prognostische waarde van de MRI bevindingen ten op zichtte van bekende prognostische factoren voor herstel. Multivariate logistische regressieanalyse werd uitgevoerd in 3 stappen: er werd een voorspellend model gemaakt met de kenmerken van de patiënt en zijn/haar rugpijn, met alleen de gerapporteerde MRI bevindingen, en met zowel de gerapporteerde MRI bevindingen als de kenmerken van de patiënt en zijn/haar pijn. Bij de 12 maanden follow-up gaf 53% van de patiënten aan hersteld te zijn. Lagere leeftijd, een betere houding/opvatting over rugpijn, acute rugpijn, neurologische symptomen van het been, en de aanwezigheid van niet-continuë rugpijn waren significant geassocieerd met herstel: AUC 0,77. Toevoeging van de MRI bevindingen leidde niet tot een betere voorspelling van herstel. In een klinisch perspectief roepen deze resultaten een aantal vragen op over het nut van de lumbale MRI in de huisartspraktijk. Maar uiteindelijk kan slechts een goed uitgevoerde gerandomiseerde trial ons meer informatie geven over de diagnostische waarde en kosteneffectiviteit.

Het gebruik van MRI door huisartsen voor de diagnose en behandeling van patiënten met lage rugpijn zou kunnen leiden tot minder onnodige verwijzingen naar de tweede lijn. Echter, gegevens omtrent de patronen van zorg na een lumbale MRI in de eerste lijn zijn schaars. Een beter begrip van deze patronen zou informatie kunnen verschaffen over hoe de uitslagen van de MRI worden gebruikt door huisartsen. Identificatie van mogelijke prognostische factoren die een verwijzing naar een specialist of een operatie voorspellen kunnen belangrijk zijn voor de patiënt en de arts. In **hoofdstuk 4** onderzochten we de relatie tussen de kenmerken van de patiënt, de kenmerken van de rugklachten, en de MRI afwijkingen met een verwijzing naar een specialist en/of een operatie aan de rug in dezelfde groep patiënten verwezen voor MRI door hun huisarts. Van de 683 geïncludeerde patiënten, rapporteerde 301 (55%) een bezoek aan een specialist in de eerste 3 maanden follow-up, en 124 (18%) ondergingen chirurgie van de wervelkolom gedurende de 12 maanden follow-up. Patiënten van een jonge leeftijd, met uitstralende pijn in het been onder de knie, met een

ernstige beperking, met in de voorgeschiedenis een rugoperatie, en/of met zenuwwortel compressie of stenose op MRI hadden meer kans om tijdens de follow-up chirurgie aan de rug te ondergaan (AUC 0,75).

Lumbale spinale stenose

Eén van de anatomische afwijkingen die gevonden kan worden met MRI is lumbale stenose. Lumbale stenose wordt vaak gebruikt om patiënten met symptomen van een anatomische vernauwing van het lumbale wervelkanaal te beschrijven. De mate van vernauwing van het wervelkanaal correleert slecht met de ernst van de symptomen en een radiologisch significante stenose kan ook worden gezien bij asymptomatische individuen. Om een definitieve diagnose te kunnen stellen zijn er dus specifieke symptomen en fysieke bevindingen gecombineerd met anatomische afwijkingen nodig. Een breed scala aan klinische, elektrodiagnostische, en radiologische testen worden momenteel gebruikt voor het diagnosticeren van lumbale spinale stenose. Het is belangrijk om de diagnostische waarde van deze testen te weten omdat vals-positieve testresultaten kunnen leiden tot onnodige operaties en/of dure of invasieve aanvullende diagnostische ingrepen.

In **hoofdstuk 5** hebben we onze vorige systematische review over de diagnostische nauwkeurigheid van de tests die worden gebruikt om lumbale spinale stenose te diagnosticeren bijgewerkt. Onze vorige systematische review bevatte studies tot maart 2004; deze nieuwe review is bijgewerkt tot maart 2011. Er voldeden 22 extra artikelen aan de inclusiecriteria, in aanvulling op de 24 opgenomen in de vorige review. Dit resulteerde in 20 artikelen over beeldvormende onderzoeken, 11 artikelen over elektrodiagnostische testen, en 15 artikelen over klinische testen. Momenteel is MRI de meest veelbelovende test voor onderzoek naar lumbale stenose, waarbij myelografie moet worden vermeden vanwege de invasiviteit en het ontbreken van een betere nauwkeurigheid. Elektrodiagnostische testen toonden geen superieure nauwkeurigheid in vergelijking met MRI. De uitslag van deze testen moet wel worden geïnterpreteerd samen met de symptomen met de hoogste diagnostische waarde, namelijk uitstralende pijn in de benen die verergerd tijdens staan, de afwezigheid van pijn bij het zitten, de verbetering van de symptomen bij het buigen naar voren, en een breed looppatroon.

Lumbale discus degeneratie

Nadat bij een patiënt met lage rugpijn ischias of een ernstige oorzaak (zoals een fractuur en maligniteit) is uitgesloten, wordt bij ongeveer 85% uiteindelijk de diagnose aspecifieke rugpijn gesteld. Bij deze patiënten worden vaak symptomatische behandelingen voorgeschreven, zoals het advies om actief te blijven en bedrust te vermijden, het voorschrijven van pijnstillende medicatie en het verwijzen naar een fysiotherapeut. Alhoewel deze aanpak relatief eenvoudig is, werkt deze helaas niet altijd goed. De beperking van deze huidige benadering wordt geïllustreerd door de vele systematische reviews die onthullen dat de bestaande behandelingen voor aspecifieke lage rugpijn in het beste geval slechts een klein effect hebben. Een reden hiervoor kan zijn dat deze 'one-size fits all' benadering ervoor zorgt dat behandelingen niet terecht komen bij die patiënten die er het meest van zouden kunnen profiteren, waardoor hun potentiële werking wordt verdunt. Het identificeren van subgroepen van patiënten voor wie bepaalde behandelingen superieur zijn is aangeduid als de 'heilige graal' van het lage rugpijn onderzoek. Een van de mogelijke subgroepen op basis van pathologisch anatomische bevindingen zijn patiënten met ernstige degeneratie van de wervelkolom. In het tweede deel van dit proefschrift voerden wij een epidemiologische studie uit naar de kenmerken van lumbale discus degeneratie.

Lumbale degeneratie wordt radiologisch gekenmerkt door de aanwezigheid van osteofyten, sclerose, en tussenwervelschijf vernauwing. Het verband tussen lage rugpijn en degeneratieve veranderingen in de lumbale wervelkolom is complex. Het is bekend dat lumbale degeneratie een mogelijke risicofactor is voor rugpijn bij volwassenen, met odds ratio's variërend van 1,3 tot 3,2. Echter, als gevolg van verschillen in de gebruikte definities voor lumbale degeneratie, zijn de studies over de relatie tussen lage rugpijn en lumbale degeneratie moeilijk met elkaar te vergelijken. Hierdoor is het nog onbekend hoe een klinisch relevante definitie voor lumbale degeneratie te definiëren.
In **hoofdstuk 6** verkenden we de associatie tussen de verschillende individuele radiografische kenmerken van lumbale degeneratie en zelfgerapporteerde lage rugpijn. In een open cohortonderzoek (the Rotterdam Study), werden 2819 lumbale röntgenfoto's beoordeeld op de aanwezigheid en de ernst van osteofyten en tussenwervelschijf vernauwing. Logistische regressie werd gebruikt om de associatie tussen verschillende definities

van de lumbale degeneratie en lage rugpijn te bepalen. Osteofyten werden het meest frequent waargenomen, en kwamen vaker voor bij mannen. Tussenwervelschijf vernauwing kwam vaker voor bij vrouwen. Beide radiografische kenmerken namen in frequentie toe met de leeftijd. Tussenwervelschijf vernauwing bleek sterker geassocieerd te zijn met lage rugpijn dan osteofyten, vooral bij mannen (OR = 1,9; 95% CI: 1,4-2,8). Tussenwervelschijf vernauwing op 2 of meer niveaus bleek sterker geassocieerd met lage rugpijn dan tussenwervelschijf vernauwing op slechts 1 niveau (OR = 2,4; 95% CI: 1,6-3,4). Na uitsluiting van niveau L5-S1 namen bijna alle associaties toe. De gevonden associaties zijn bijna even groot als de associaties tussen kniepijn en radiologische knieartrose en zelfs iets groter dan de associatie tussen handpijn en radiologische handartrose in dezelfde populatie. Uit onze data kan een nuttige definitie voor lumbale degeneratie worden afgeleid; tussenwervelschijf vernauwing op 2 of meer niveaus van L1/2 tot L4/5. Deze definitie kan een klinisch relevant fenotype zijn voor toekomstig genetisch en epidemiologisch onderzoek naar lumbale degeneratie.

Hoewel lumbale degeneratie niet kan worden gedefinieerd als artrose omdat alleen de facetgewrichten van de rug synovium bevatten, wordt de term toch vaak gebruikt als een indicatie voor artrose van de wervelkolom. Klinische artrose van de knie en heup omvat naast pijn ook ochtendstijfheid (ACR). In **hoofdstuk 7** verkenden we daarom de associatie tussen ochtendstijfheid van de rug en lumbale degeneratie. Gegevens uit de studie van hoofdstuk 6 werden gebruikt om de associatie tussen ochtendstijfheid van de rug en twee definities voor lumbale degeneratie (‘osteofyten’ en ‘tussenwervelschijf vernauwing’) te onderzoeken. Tevens werd de associatie tussen ochtendstijfheid gecombineerd met lage rugpijn en lumbale degeneratie onderzocht. Vergelijkbare analyses werden uitgevoerd met knie en heup pijn, ochtendstijfheid in de benen en radiografische knie en heup artrose. Beide definities voor lumbale degeneratie waren geassocieerd met ochtendstijfheid van de rug (OR = 1,3; 95% CI: 1,1-1,6 voor ‘osteofyten’ en OR = 1,8; 95% CI: 1,4-2,2 voor ‘tussenwervelschijf vernauwing’). Beide odds ratio’s stegen wanneer ochtendstijfheid werd gecombineerd met lage rugpijn (OR = 1,5; 95% CI: 1,1-2,0 voor ‘osteofyten’ en OR = 2,5; 95% CI: 1,9-3,4 voor ‘tussenwervelschijf vernauwing’). De omvang van de associatie tussen ‘tussenwervelschijf vernauwing’ en ochtendstijfheid van de rug is vergelijkbaar met de associatie tussen ochtendstijfheid in de benen en knie-artrose (OR = 3,0; 95% CI: 2,1-4,1). We kunnen hieruit concluderen

dat klachten van ochtendstijfheid van de rug voor artsen een aanwijzing kan zijn voor lumbale degeneratie.

In **hoofdstuk 8** hebben we gekeken naar de mogelijke associatie tussen artrose van de wervelkolom en heuppijn. Het differentiëren tussen rug- en heuppijn bij patiënten met klassieke symptomen is meestal niet moeilijk. Echter, in sommige gevallen, presenteren patiënten zich met aspecifieke pijnklachten in de bil, de laterale heup of dij. De differentiële diagnose van deze aspecifieke pijn is breed en omvat uitstralende pijn vanuit de lumbale wervelkolom. De differentiatie tussen symptomen wijzend op een heup aandoening versus een aandoening van de rug is van belang bij het geven van een behandeling, vooral als de behandeling een grote reconstructieve operatie zoals een heupprothese inhoud. Preoperatieve identificatie van de factoren die samenhangen met heuppijn als gevolg van een probleem in de rug zou artsen kunnen helpen bij het identificeren van die groep patiënten die geen volledige verlichting van de pijn zullen ervaren bij een heupprothese.

In **hoofdstuk 8** hebben we dan ook gekeken naar de associatie van zelf-gerapporteerde heuppijn met de verschillende individuele radiografische kenmerken van lumbale degeneratie in de 2819 gescoorde röntgenfoto's van de Rotterdam-studie (hoofdstuk 6). De aanwezigheid van tussenwervelschijf vernauwing op het niveau L1/L2 was significant geassocieerd met heuppijn (mannen OR = 2,0; 95% CI: 1,1-3,8 en vrouwen OR = 1,7; 95% CI: 1,1-2,5). De aanwezigheid van tussenwervelschijf vernauwing op het niveau L2/L3 was alleen significant geassocieerd met pijn in de heup bij vrouwen. De sterkte van de associaties nam toe bij chronische heuppijn, vooral bij mannen (L1/L2 OR = 2,5; 95% Cl: 1,3-5,0). De aanwezigheid van tussenwervelschijf vernauwing op de lagere niveaus (L3/L4/L5/S1) was niet significant geassocieerd met pijn in de heup. We concludeerden dat in geval van onzekerheid over de oorzaak van de heuppijn, de evaluatie van een lumbale röntgenfoto kan helpen die patiënten te identificeren die mogelijk pijn hebben als gevolg van slijtage van de lumbale wervelkolom. Echter, er zijn goed opgezette studies nodig om deze hypothese te controleren.

Afsluitend reflecteerden we in **hoofdstuk 9** op de belangrijkste bevindingen van dit proefschrift en beschreven we de hieruit volgende implicaties voor de klinische praktijk en onderzoek.

Curriculum vitae

Evelien de Schepper is geboren op 16 januari 1983 te Bergen op Zoom. Zij is de dochter van Fred de Schepper en Mia Famaey. Zij groeide op samen met haar jongere zus Rosanne in Bergen op Zoom.

Na het behalen van haar vwo-diploma aan het R.K. Gymnasium Juvenaat 't Hart, begon zij in 2001 aan de Universiteit van Gent in België aan de studie geneeskunde. In 2002 verruilde zij deze universiteit voor de Erasmus Universiteit in Rotterdam waar zij haar studie geneeskunde voortzette. Zes jaar later rondde zij deze succesvol af. Na een korte klinische ervaring op de eerste hulp van het Vlietland Ziekenhuis in Schiedam startte zij in 2009 als AIOTHO (arts in opleiding tot huisarts en onderzoeker) op de afdeling Huisartsgeneeskunde van het Erasmus MC. Zij werkte als promovendi op het project naar diagnostiek bij lage rugpijnklachten onder begeleiding van prof. Sita Bierma-Zeinstra, prof. Bart Koes en copromotor Pim Luijsterburg. In 2011 behaalde zij een Master of Science in Clinical Epidemiology aan het Netherlands Institute for Health Sciences (NIHES).

Begin 2015 heeft zij de huisartsopleiding afgerond en is zij kortdurend werkzaam geweest als HIDHA (huisarts in dienst van huisarts). Sindsdien werkt zij als onderzoeker en docent op de afdeling Huisartsgeneeskunde van het Erasmus MC.

Evelien woont samen met Andra Veraart en zij hebben samen een dochter Lisa (geboren 8 mei 2014).

List of publications

This thesis

de Schepper EIT, Damen J, van Meurs JB, Ginai AZ, Popham M, Hofman A, Koes BW, Bierma-Zeinstra SMA. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. Spine (Phila Pa 1976). 2010 Mar 1;35(5):531-6.

Scheele J, *de Schepper EIT*, van Meurs JB, Hofman A, Koes BW, Luijsterburg PA, Bierma-Zeinstra SMA. Association between spinal morning stiffness and lumbar disc degeneration: the Rotterdam Study. Osteoarthritis Cartilage. 2012 Sep;20(9):982-7.

de Schepper EIT, Damen J, Bos PK, Hofman A, Koes BW, Bierma-Zeinstra SMA. Disk degeneration of the upper lumbar disks is associated with hip pain. Eur Spine J. 2013 Apr;22(4):721-6.

de Schepper EIT, Overdevest GM, Suri P, Peul WC, Oei EH, Koes BW, Bierma-Zeinstra SMA, Luijsterburg PA. Diagnosis of lumbar spinal stenosis: an updated systematic review of the accuracy of diagnostic tests. Spine (Phila Pa 1976). 2013 Apr 15;38(8):E469-81.

de Schepper EIT, Bart W. Koes, Erik F.H. Veldhuizen, Edwin H.G. Oei, Sita M.A. Bierma-Zeinstra, Pim A.J. Luijsterburg. Prevalence of Spinal Pathology in Back Pain Patients Referred for MRI in Primary Care. Family Practice 2015: Accepted for publication

de Schepper EIT, Bart W. Koes, Edwin H.G. Oei, Sita M.A. Bierma-Zeinstra, Pim A.J. Luijsterburg. Does MRI Add to the Prediction of Recovery in Low Back Pain Patients in General Practice? Eur Spine J. 2015: Revision submitted

de Schepper EIT, Bart W. Koes, Erik F.H. Veldhuizen, Edwin H.G. Oei, Sita M.A. Bierma-Zeinstra, Pim A.J. Luijsterburg. Healthcare Service Use After Lumbar Spine MRI in General Practice. Spine (Phila Pa 1976) 2015: Under revision

Other publications

Williams FM, Popham M, Hart DJ, *de Schepper EIT*, Bierma-Zeinstra SMA, Hofman A, Uitterlinden AG, Arden NK, Cooper C, Spector TD, Valdes AM, van Meurs J. GDF5 single-nucleotide polymorphism rs143383 is associated with lumbar disc degeneration in Northern European women. Arthritis Rheum. 2011 Mar;63(3):708-12.

Castaño-Betancourt MC, Oei L, Rivadeneira F, *de Schepper EIT*, Hofman A, Bierma-Zeinstra S, Pols HA, Uitterlinden AG, Van Meurs JB. Association of lumbar disc degeneration with osteoporotic fractures; the Rotterdam study and meta-analysis from systematic review. Bone. 2013 Nov;57(1):284-9.

Evelien de Schepper. MRI-diagnostiek bij rugpijn. Huisarts en wetenschap 03/2014; 57(3):136-136.

Zwinkels ML, Hassan H, *de Schepper EIT*. Antihypertensive medication and the effects of black tea on blood pressure variation. Am J Clin Nutr. 2013 Sep;98(3):857

PhD Portfolio

PhD training	Year	ECTS
Courses		
Master of Science in Clinical Epidemiology, NIHES, Rotterdam	2009-2014	70
Professional education		
Vocational training for general practitioner, Erasmus MC, Rotterdam	2010-2015	
Conferences		
OARSI World Congress, Philadelphia	2013	1
Oral presentations		
OARSI World Congress, Brussels	2010	2
World Congress on Low Back and Pelvic Girdle Pain, Dubai	2013	2
NAPCRG Annual Meeting, New York	2014	2
NHG Wetenschapsdag, Rotterdam	2015	2
Poster Presentations		
Back Pain Forum, Boston	2009	1
EULAR Annual Congress, Copenhagen	2009	1
NHG Wetenschapsdag, Utrecht	2009	1
Primus, Rotterdam	2010	1
NHG Congres, Groningen	2010	1
Back Pain Forum, Melbourne	2011	1
Back Pain Forum, Odense	2012	1
OARSI World Congress, Barcelona	2012	1
World Congress on Low Back and Pelvic Girdle Pain, Dubai	2013	1
OARSI World Congress, Seattle	2015	1
Teaching activities		
Supervising student session 'How to judge a paper'	2013	1
Workshop for GP-trainees and GP-tutors: Diagnostics in low back pain, STarT Back Screening Tool	2013	1
Developing e-learning program for medical students: Clinical Reasoning, Erasmus MC	2015	6

You've always got my back & I sure do love you for that

Andra & Lisa, mam & pap, Rosanne & Paul, Jac & Clara, Joska, Nikita & Rinske, Kioga, Tuyet & Terry, Oma, al mijn neven, nichten, ooms & tantes, Henrike & Laura, Sita, Pim, Bart, Patrick, Lex, Herman, Prof.dr. Ir. A. Burdorf, Prof.dr. G.P. Krestin, Prof.dr. W.C. Peul, Prof.dr. J.A.N. Verhaar, Prof.dr. J. Gussekloo, Pradeep Suri, Edwin Oei, Erik Veldhuizen, Gijsbert Overdevest, Joyce van Meurs, Abida Ginai, Maria Popham, Albert Hofman, Jantine Scheele, Pieter Bos, patiënten ERGO, patiënten & collega's MRI Centrum, Coen & Mase Sutterland, Rutger Boggia, Johan & Coby Eckhardt, Frans van den Brule & andere Waddinxveen collega's, Ignace Ong, Hans van der Sijde, Erik Paling, Joyce Landvreugd, Aafke, Adinda, Alex & Marly, Alyt, Annemieke, Arianne, Arthur, Carolien, David, Desiree, Diana, Dieuwke, Erwin, Fiona, Gijs, Jacoline, Jan, Jasper, Joost, Jorien, Jos, Josje, Karin, Kelly, Laraine, Leo, Lonny & Ankie, Manuel, Marieke, Marienke, Mariet, Marijke, Marlies, Mary, Mathijs, Metthilde, Nadine, Nynke, Nienke, Pauline, Pepijn, René, Rianne, Rianne, Roxanne, Saskia, Theun, Toke, Wendy, Winifred, Australian low back pain researchers, Groepsgenootjes jaar 1,2 & 3, Hilde, Anne, Frits, Rene, Mieke, Eerstehulp buddies Ruwaard van Putten, Marijn, Journal Club Daniel, Jurgen, Marjolein & Stephan, Honkbalmeiden incl. Ritchie, Carlo & Jan, de rest van mijn Neptunus familie incl. Naat, Stees & de Stuijtjes, Softbalteam Wizards of BoZ, Rene & Margreet, Kaj & Christine, Nieke, Henny, Liesbeth & Marcel, Jan Hendrik, Wahida, Barbara, Noortje, Nicola, Jeroen, Seb, Yrjo, Nils & Marlou

Diagnostics in low back pain

The aim of the work presented in this thesis is to gain more insight into the diagnostic tools used in patients with low back pain.

In recent years, general practitioners in the Netherlands can refer low back pain patients for MRI of the lumbar spine themselves. However, data on the use of MRI by general practitioners in the Netherlands are still lacking. In the first part of this thesis, we examined how lumbar MRI is currently used by general practitioners in the Netherlands.

Identifying subgroups of patients for whom different treatments are superior has been referred to as the 'Holy Grail' of low back pain research. One of the possible subgroups based on pathoanatomical findings are patients with degeneration of the spine. In the second part of this thesis, we performed an epidemiological study on the characteristics of lumbar disc degeneration.

Stellingen

Diagnostics in low back pain

1. Patiënten met lage rugpijn die worden verwezen door hun huisarts voor een MRI hebben vaak ernstige en/of langdurige rugklachten (dit proefschrift)
2. De uitslag van een lumbale MRI scan aangevraagd door de huisarts heeft geen toegevoegde prognostische waarde ten opzichte van informatie uit de anamnese (dit proefschrift)
3. Patiënten met positieve overtuigingen over hun rugpijn hebben een grotere kans op herstel (dit proefschrift)
4. Lumbale discusdegeneratie is gerelateerd aan het hebben van rugpijn; deze relatie is sterker wanneer meerdere niveaus zijn aangedaan (dit proefschrift)
5. Ochtendstijfheid van de rug kan een aanwijzing zijn voor lumbale discusdegeneratie (dit proefschrift)
6. Een diagnose kan van waarde zijn, zelfs zonder de beschikbaarheid van een effectieve behandeling, omdat het een logische route kan bieden voor de ontwikkeling en het testen van toekomstige interventies – Hancock et al. 2011
7. Het vervangen van een gewricht door een prothese resulteert lang niet altijd in een gelukkige patiënt – Beswick et al. 2012
8. Educatie bij aspecifieke rugpijn door de huisarts zelf zorgt voor een betere geruststelling dan wanneer deze gegeven wordt door een andere zorgverlener uit de eerste lijn zoals een doktersassistente of fysiotherapeut – Traeger et al. 2015
9. Artsen zullen zich moeten instellen op patiënten met zelf gegenereerde gezondheidsinformatie; zij krijgen een belangrijke rol bij het geven van betekenis aan data – Hengst et al. 2014
10. E-learning met behulp van virtuele patiënten is een uitstekende manier om de ontwikkeling van het klinisch redeneren bij studenten te ondersteunen – Bateman et al. 2013
11. Een theorie is pas compleet als je hem zo helder hebt dat je hem aan de eerste de beste voorbijganger kunt uitleggen – David Hilbert 1900