

What diagnostic tests are useful for low back pain?

Jon D. Lurie* MD, MS

Assistant Professor of Medicine and of Community and Family Medicine
Dartmouth Medical School and Dartmouth-Hitchcock Medical Center, NH, USA

Low back pain (LBP) is a common problem that poses some interesting and difficult diagnostic problems. It is typically benign and self-limited, but it is occasionally the presenting symptom of serious systemic disease. The general diagnostic approach to low back pain is to check for 'red flags' in the history and physical that suggest the presence of malignancy, infection or spondyloarthritides, and for neurological compromise that could indicate that surgery is required (cauda equina syndrome) or may be beneficial (such as herniated discs or spinal stenosis that have not improved with conservative care). In the absence of these features, imaging is of limited value. Recent research has begun to evaluate subgroups with 'non-specific' low back pain that seem to benefit from specific interventions such as median branch or sacroiliac joint injections, manipulation, or specific exercises, but these require further investigation and validation.

Key words: clinical examination; diagnosis; imaging; low back pain; sensitivity; specificity; spinal disorders.

Although low back pain (LBP) is a common problem, it poses some interesting and difficult diagnostic problems. Typically, it is typically benign and self-limited although it is occasionally the presenting symptom of serious systemic disease, such as cancer or infection, or of a surgical emergency, such as with cauda equina compression. The major diagnostic task is to distinguish the $\geq 95\%$ who have a benign musculoskeletal pain syndrome from the small minority with a serious, specific disease process that requires timely and specific therapy. This process needs to be thoughtfully targeted to avoid overtesting those with benign musculoskeletal pain while effectively identifying serious diseases. The major components of the differential diagnosis of LBP are as follows (these have been adapted Refs. 2,5,9,15.

* Address: SPORT/The Spine Center, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756, USA.
Tel.: +1 603 653 3575; Fax: +1 603 653 9502.

E-mail address: jon.d.lurie@dartmouth.edu

- Regional mechanical low back pain ($\pm 90\%$):
 - non-specific mechanical low back pain (sprain, strain, lumbago, etc.)
 - degenerative changes in discs and/or facet joints
 - osteoporotic compression fractures
 - traumatic fractures
 - deformity (severe scoliosis, kyphosis)
 - symptomatic isthmic spondylolisthesis
- Mechanical low back pain with neurogenic leg pain (7–10%):
 - intervertebral disc herniation
 - spinal stenosis
 - spinal stenosis associated with degenerative spondylolisthesis
- Non-mechanical spine disorders ($\pm 1\%$):
 - neoplasia (metastases, lymphoid tumors, spinal cord tumors, etc.)
 - infection (infective spondylitis, epidural abscess, endocarditis, herpes zoster, Lyme disease)
 - seronegative spondyloarthritides (ankylosing spondylitis, psoriatic arthritis, reactive arthritis, Reiter's syndrome, inflammatory bowel disease)
- Visceral disease (1–2%):
 - pelvic (prostatitis, endometriosis, pelvic inflammatory disease)
 - renal (nephrolithiasis, pyelonephritis, renal papillary necrosis)
 - aortic aneurysm
 - gastrointestinal (pancreatitis, cholecystitis, peptic ulcer disease)
- Miscellaneous:
 - Paget's disease
 - parathyroid disease
 - hemoglobinopathies

There is a great desire to improve our ability to identify subgroups of patients within the broader group of 'non-specific, musculoskeletal LBP' who might improve with specific targeted therapies. The enthusiasm of the physician, and the patient, for this diagnostic task, however, needs to be tempered by a realistic assessment of the limitation of pathoanatomic diagnosis in the complex, multifactorial, biopsychosocial context of LBP illness and the dangers of 'labeling' and 'over-medicalizing' patients with otherwise benign backache.^{1–6}

This chapter attempts to answer the following questions: (i) What is the relative value of clinical history in LBP? (ii) What is the relative value of physical examination in LBP? (iii) What is the relative value of imaging in LBP? For each question we will focus on the following diagnostic tasks: (a) Is there a serious underlying systemic disease causing the LBP? (b) Is there neurological impairment that might require, or at least benefit from, surgical evaluation? (c) If the answer to the first two questions is no, does the patient fall into a subgroup which is more or less likely to benefit from some specific targeted therapy?

WHAT IS THE RELATIVE VALUE OF CLINICAL HISTORY IN LOW BACK PAIN?

A careful clinical history is the starting point for any evaluation of a patient with LBP. It not only supplies key information for the diagnosis but provides an important first

step towards the rapport that is crucial for ongoing management.⁷ Patients with LBP might not feel listened to and might not feel that their pain has been validated, and this can strongly affect their satisfaction with care. In one comparison of patients' perceptions of care delivered by family physicians and chiropractors, patients were much more satisfied with the care delivered by chiropractors; patients of chiropractors felt more satisfied with the amount of time spent listening to their description of the pain, felt the provider believed their pain was real, felt the provider understood their concerns about the cause of pain, and rated the provider as more confident in the care they delivered.⁸ This therapeutic alliance might be far more important to how the patients ultimately do than the specific diagnoses made or treatments being delivered.⁸

IS THERE A SERIOUS UNDERLYING SYSTEMIC DISEASE CAUSING THE BACK PAIN?

When evaluating a patient with LBP, the initial task is to quickly assess the likelihood of a serious underlying systemic disease such as malignancy. Malignancy is estimated to account for about 0.7% of cases seeking care for LBP.⁵ History is more useful than the physical examination for early detection of the possibility of malignancy.⁹ The classic 'red flags' for malignancy include age > 50, previous history of cancer, unexplained weight loss (> 4.5 kg over 6 months), failure to improve after 1 month of therapy, and no relief with bedrest. The only finding specific enough to significantly increase the odds of malignancy is a previous history of cancer with a likelihood ratio (LR) of 15. When present, the other red flags have LRs of between 2 and 3, which raise the odds only modestly.⁹ Pain that is worse at night or with recumbency, particularly when patients sleep in a chair to avoid pain, is very worrisome for malignancy or infection, although the precise sensitivity and specificity are unknown.^{10,11} The most sensitive 'red flag' is no relief with bedrest (sensitivity > 0.9). Absence of this indicator, i.e. pain is relieved with lying down, significantly reduces the odds of malignancy to about a fifth of the starting odds (LR 0.21). However, this finding is rather non-specific and the positive LR is only 1.7; most patients who report a lack of relief with recumbency/rest will still have benign backache.⁹ The diagnostic utility of sciatica, defined as pain in the distribution of a lumbar nerve root radiating into the leg below the knee and often accompanied by neurosensory symptoms or deficits¹², as an indicator of malignancy is unclear. A systematic review found one study reporting a sensitivity of 0.93 and a specificity of 0.78, suggesting that sciatica, when present, raised the odds of malignancy by 4-fold. Another study, however, found a sensitivity of only 0.58.¹³

Because of the seriousness of malignancy, one would like as sensitive a test as possible so as to be able to confidently rule out this possibility; no single historical feature adequately fits the bill. Using a combination of features (age > 50, history of cancer, unexplained weight loss, or failure of conservative therapy) obtained 100% sensitivity; the absence of all four confidently rules out malignancy.⁹ This combination of red flags was designed to be very sensitive but because of the inherent trade-off between sensitivity and specificity it suffers with a specificity of only 0.6. Further work-up is often recommended for patients with one of the 'red flags' but the likelihood of malignancy remains low, at 1–10%.^{11,14}

Infective spondylitis is a very rare cause of LBP and is estimated to account for 0.01% of cases.¹⁵ The incidence increases with age, with the majority of cases in patients > 50

years of age.¹⁶ Intravenous drug use, urinary tract infection, indwelling urinary catheters and skin infections are known risk factors with a sensitivity of about 0.4; specificity of these findings is unknown. Fever is a feature that should strongly suggest infection; the specificity of fever for infection as a cause of LBP is estimated to be about 98% with a LR of about 25.^{11,17} Unfortunately, fever has a sensitivity of only about 50% and thus, unlike the red flags for malignancy, although the presence of fever should raise strong concern for infective spondylitis, absence of fever does not significantly lower the odds of infection. The clinician must maintain a high index of suspicion if other unusual features are present.¹¹

With an aging population and better treatment options for osteoporosis, it is becoming more important to recognize compression fractures as a cause of LBP. Age, history of trauma, and history of steroid use are the historical features and have been well studied as predictors in LBP. Compression fractures are estimated to represent about 4% of LBP cases.¹⁵ Being less than 50 years of age significantly lowers the odds of compression fracture, with an LR of 0.26, whereas being over 70 increases the odds of compression fracture with a LR of 5.5.⁹ Because many osteoporotic fractures occur in the absence of trauma, a history of trauma is not particularly useful; with a sensitivity of 0.3 and a specificity of 0.85 its presence or absence does not markedly alter the odds of compression fracture.⁹ Finally, corticosteroid use has been found to be fairly specific, with an LR of 12; compression fracture needs to be strongly considered in any LBP patient using corticosteroids.⁹

The seronegative spondyloarthritis need to be considered when progressive back pain is present in a young patient with inflammatory symptoms, such as morning stiffness and improvement with exercise.¹⁵ Age of onset < 40 years of age was 100% sensitive, and although, therefore, excellent for ruling out a diagnosis of ankylosing spondylitis (AS), was extremely non-specific.⁹ In a recent review, the most specific historical feature was four out of five positive responses to the following battery of questions: (i) Did the back discomfort begin before age 40? (ii) Did the discomfort begin slowly? (iii) Has the discomfort persisted for > 3 months? (iv) Was morning stiffness a problem? (v) Did the discomfort improve with exercise? Although the specificity was reasonable at 0.82, the sensitivity was low at 0.23, giving an LR for four out of five positive responses of only 1.3. However, in a study comparing early AS with lumbar disc disease patients, Sadowska-Wroblewska et al found several highly specific historical features, including a history of iritis (100% specific), family history of AS (100% specific), thoracic stiffness (100% specific), thoracic pain (97% specific, LR 1.3) and heel pain (90% specific, LR 1.6).¹⁸ This study, however, did not have a clear gold standard for the diagnosis of early AS and the authors did not report on the reliability of their history taking. In another study, the apparently straightforward history item of 'up during the night because of pain' had poor reliability—17% of patients recording this symptoms later denied ever having had this problem.¹⁹

IS THERE NEUROLOGICAL IMPAIRMENT THAT MIGHT REQUIRE, OR BENEFIT FROM, SURGICAL EVALUATION?

Cauda equina and spinal cord compression syndrome are the most important neurological entities in the differential diagnosis because they represent surgical emergencies. Cord compression can occur in the setting of spinal tumors or epidural abscesses, or with massive midline intervertebral disc herniation (IDH).

Fortunately, this entity is quite rare, accounting for an estimated 0.04% of LBP cases. Unilateral or bilateral leg pain, numbness, and/or weakness are common, each occurring in over 80% of cases.¹⁷ The most sensitive finding seems to be urinary retention, which has a sensitivity of 0.9 and a specificity of 0.95; the LR of urinary retention is thus 18, whereas the LR for the absence of urinary retention is 0.1.¹⁷ Later, urination retention can give way to frank incontinence, which is an extremely worrisome symptom. Care must be taken in the history as patients with sciatica can report difficulties with toileting that, on closer questioning, is really pain with sitting on the toilet or increased pain with Valsalva during toileting. These should not be confused with urinary retention or incontinence as worrisome features for cauda equina syndrome.

The typical history suggesting IDH is sciatica, which is often worse with coughing, sneezing, or the Valsalva maneuver.²⁰ One review reported that typical sciatica had an excellent performance for diagnosing disc herniation with a sensitivity of 0.95 and a specificity of 0.88.⁹ However, others found much more variable results in the literature, with sensitivity ranging from 0.8 to 0.99 and specificity from 0.06 to 0.15.^{21,22} This discrepancy might be at least partially due to limited reliability. Assessment of the interobserver reliability of a 'typical' radicular pain pattern yielded a kappa value of only 0.24; the assessment of increased pain with coughing/sneezing/straining had greater reliability, with a kappa of 0.64, and the reliability of a report of paresthesias was intermediate at kappa = 0.47.²³ A study using selective lumbar spinal nerve blocks showed significant variability between individuals in the resulting areas of numbness, with deviation from the classical dermatomes.²⁴ This underlying variability can explain some of the inconsistency in assessing pain radiation patterns. Other historical features for disc herniation have been paid very little attention in the literature and their diagnostic value remains unclear.²²

Symptoms suggestive of spinal stenosis (SpS) tend to occur in older patients, may be bilateral and are most typically leg pain when standing or walking that resolves with sitting down or leaning over; pain does not improve with resting upright and is often worse when walking downhill than uphill.¹⁵ Age > 65 is modestly sensitive and specific for a diagnosis of SpS, with a reported sensitivity of 0.77 and specificity 0.69.⁹ The absence of pain when seated has low sensitivity (0.46) but this is very specific (0.93), thus the report of complete relief with sitting, when present, increases the odds of SpS by a factor of 6.6.⁹ The improvement, rather than the absence, of pain with sitting was somewhat more sensitive (0.52) but is significantly less specific (0.83); symptoms that are worse with walking had essentially no diagnostic value, with a positive LR of 1.0 and a negative LR of 0.97.⁹

DOES THE PATIENT FALL INTO A SUBGROUP THAT IS MORE OR LESS LIKELY TO BENEFIT FROM SOME SPECIFIC TARGETED THERAPY?

The effectiveness of most treatments for low back pain are either disappointing or controversial.²⁵ One potential reason is the lack of a coherent and reliable classification system which allow for the selection of subgroups who might respond better (or worse) to specific treatments. Lumping together the heterogeneous group of all LBP patients provides the potential to wash-out real effects that might appear if well selected subgroups had been tested.²⁶ The identification of well-defined subgroups within LBP has been identified as a key research priority.²⁷ Recently, some studies have

begun to make progress in this area. Most use a combination of history and exam to put patients into groups and do not clearly distinguish between the components of the predictive model. As a result, we will discuss those using significant components of history and patient self-report in this section and those relying more heavily on physical exam in the next section.

Although the facet joint has been identified as a potential pain generator in LBP for over 70 years, there is still significant controversy and uncertainty regarding the prevalence and clinical features of this entity.²⁸ The prevalence of pain responsive to placebo-controlled, anesthetic injections of the facets has been reported to be somewhere between 10 and 40%, although these have generally been in preselected populations.^{28,29} Two recent, well-designed studies have yielded some important insights into the characteristics of facet injection responsive chronic LBP.

Perhaps the most striking finding regarding facet injections is the high placebo response rate: between 20 and 40% of patients in these studies responded to placebo injections, which should give one pause for thought when interpreting the results.^{28,29} Schwarzer et al found no significant differences in any characteristics of demographics, history, or physical (including age, pain intensity, exacerbation with motion, lumbar range of movement (ROM), straight leg raising (SLR), pain location, pain with extension \pm rotation, and psychometric testing) between those with and those without facet-injection-responsive LBP.²⁹ Revel et al did find some differences: the constellation of reporting pain well relieved by recumbency and at least four out of six other findings—age > 65 , no exacerbation with coughing, no exacerbation with forward flexion, no exacerbation when arising from flexion, no exacerbation with hyperextension, and no exacerbation with extension/rotation—had a sensitivity of 92% and specificity of 80% for identifying response to facet injection.²⁸ It is interesting to note that increased pain with extension \pm rotation was not simply unhelpful in identifying facet-injection-responsive pain; it actually helped select patients who were less likely to respond. These findings are intriguing and require further confirmation in other studies.

Recent studies have tried to evaluate predictors of manipulation responsive acute LBP. Flynn et al developed a clinical prediction rule to identify patients with LBP who were likely to show improvement with manipulation.³⁰ Duration of symptoms < 16 days and hip internal rotation (tested prone) of at least 35° had the highest LRs for predicting success with manipulation (4.4 and 3.3, respectively); other features in the prediction rule included lumbar segmental hypomobility, as judged by downward pressure on the spinous processes with the hypothenar eminence, no symptoms distal to the knee, and a score of < 19 on the work subscale of the Fear-Avoidance Beliefs Questionnaire (see http://www.annals.org/content/vol141/issue12/images/data/920/DCI/Childs_video_141-12-920-DCI-1.swf). The presence of four or five of these findings increased the odds of short-term response to manipulation 25-fold, although the confidence intervals were wide in this small study. In a validation study, a score of four or more on the rule had an LR for success at 1 week of 13.2, and a significant interaction was seen in which patients with a prediction score of four or more had better outcomes when they received manipulation compared to exercise; those who were negative on the rule had similar outcomes with either manipulation or exercise and there was no overall difference in outcome between the prediction rule positive group and the two prediction-rule-negative groups.³¹ These results are promising but need further confirmation to assess their robustness and generalizability.³²

Finally, psychological factors remain an important part of the history in patients with LBP. A recent systematic review found that psychological distress/depressive mood and

somatization were consistent and important predictors of chronicity and poor outcome in LBP.³³

WHAT IS THE RELATIVE VALUE OF PHYSICAL EXAMINATION IN LBP?

A great number of physical exam features and maneuvers have been studied and used in the evaluation of LBP. Systematic reviews have found a number of methodological problems in many of these studies.^{13,26} In addition, many examination features have been found to lack either reliability or diagnostic value but continue to be used. It is important to realize that there is a somewhat complicated relationship between reliability and diagnostic value. Tests that are unreliable cannot have reproducible diagnostic value, and the finding of limited diagnostic utility often results from the lack of reliability. Typically, these tests or maneuvers should be abandoned. However, if findings have poor or limited reliability but are found in controlled settings to be highly predictive, then efforts need to be undertaken to standardize and operationalize the finding in a way that maintains its predictive power but increases its reliability, and to better train clinicians in a reliable assessment of the finding to take advantage of its diagnostic value.

IS THERE A SERIOUS UNDERLYING SYSTEMIC DISEASE CAUSING THE BACK PAIN?

The physical examination is less helpful than the history for identifying malignancy at early stages (Table 1). Of note: The majority of spinal malignancy is metastatic from breast, lung or prostate, so these areas should be examined carefully when malignancy is suspected.¹⁷ Focal bony tenderness in the midline might be somewhat suggestive; it is a moderately reliable finding ($\kappa = 0.4$; soft-tissue or paravertebral tenderness is less reproducible with a κ of 0.24) and has sensitivity ranging from 0.15 to 0.60 and specificity of 0.60–0.78.¹³ Weakness and sensory loss have had low sensitivity around 0.40 but fairly high specificities of 0.8–0.9.¹³

As mentioned above, fever has a low sensitivity for infective spondylitis, occurring in half or fewer of cases. Percussion tenderness has reasonable sensitivity at 0.86 but poor specificity of only 0.6. Thus the absence of percussion tenderness lowers the odds of spondylitis to about 25% of the initial odds, although its presence raises the odds more modestly with an LR of 2.2.¹⁷

Again, the physical examination seems to be less helpful than the history for diagnosing spondyloarthritis, specifically AS. Tests for sacroiliac joint tenderness have modest reproducibility at best and have been found to have low sensitivity for AS (in the range 0.1–0.27).^{13,17} Schober's sign also has sensitivity of 0.30.^{13,17} Reduced lateral mobility was somewhat more helpful, with a sensitivity of 0.52 and a specificity of 0.82, raising the odds of AS by a factor of 2.9. Reduced chest expansion is highly specific at 0.99 but very insensitive at 0.09; thus when present it increases the odds of AS 9-fold, but its absence is uninformative.⁹

Table 1. Diagnostic performance of clinical examination for serious systemic causes of LBP (data from Refs. 9,13,17,18,35).

Condition	Test	Reliability (kappa)	Sensitivity	Specificity	LR when present	LR when absent
Malignancy	Prior history of cancer	?	0.31	0.98	15.5	0.7
	No relief with bed rest	?	> 90	0.46	1.7	0.21
	Any one of: age > 50, history of cancer, unexplained weight loss, or failure of conservative therapy	?	1.0	0.6	2.5	0.0
	Bony midline tenderness	0.4	0.15–0.6	0.6–0.78		
	Sensory or motor deficit	0.65–1.0	0.4	0.8–0.9	≈ 3	≈ 0.7
	Infective spondylitis	Injection drug use, urinary infection, indwelling urinary catheter, or skin infection		0.4		
Fever		?	0.5	0.98	25	0.51
Percussion tenderness		?	0.86	0.6	2.1	0.23
Compression fracture		Age > 50		0.84	0.61	2.2
	Age > 70		0.22	0.96	5.5	0.81
	Trauma	?	0.3	0.85	2.0	0.82
	Corticosteroid use	?	0.06	0.995	12	0.94
Ankylosing spondylitis	Onset < 40 years old	?	1.0	0.07	1.07	0.0
	Reduced lateral mobility	?	0.52	0.82	2.9	0.59
	Reduced chest expansion	?	0.09	0.99	9.0	0.92

IS THERE NEUROLOGICAL IMPAIRMENT THAT MIGHT REQUIRE, OR BENEFIT FROM, SURGICAL EVALUATION?

The physical examination is less sensitive than a history of urinary retention for diagnosing cauda equina syndrome but several examination features can be helpful (Table 2). Sensory and motor deficits, as well as a positive SLR test occur in about 80% of cases, although the specificity of these findings is going to be low because cauda equina syndrome is rare.¹⁷ Findings that should be somewhat more specific for cauda equina syndrome include saddle anesthesia, which has a sensitivity of about 75%, and diminished anal sphincter tone, which has a sensitivity of 60–80%; the precise specificity of these findings is unclear.¹⁷

Table 2. Diagnostic performance of clinical exam for LBP associated with neurologic impairment (data from Refs. 9,13,17,34–36,38).

Condition	Test	Reliability (kappa)	Sensitivity	Specificity	LR when present	LR when absent
Cauda equina compression	Urinary retention	?	0.9	0.95	18	0.1
	Saddle anesthesia	?	0.75	?		
	Decreased sphincter tone		0.6–0.8			
Intervertebral disc herniation	'Typical' pain pattern	0.24	0.8–0.99	0.06–0.15	≈ 1	Variable
	Worse with coughing	0.64	0.74	?		
	Straight leg raise (SLR)	0.68	0.91	0.26	1.2	0.3
	Crossed SLR	0.7	0.29	0.88	2.4	0.8
	Bell test	0.51–0.67	0.35–0.53	0.63	≈ 1	≈ 1
	Hyperextension test	0.35–0.5	0.4–0.47	0.71	≈ 1.5	≈ 0.8
Spinal stenosis	Age > 65		0.77	0.69	2.5	0.3
	No pain when seated		0.46	0.93	6.6	0.6
	Better when seated		0.52	0.83	3.1	0.6
	Symptoms worse with walking		0.71	0.3	1.0	0.97
	No pain with flexion		0.79	0.44	1.4	0.5
	Wide-based gait		0.43	0.97	14	0.6
	Abnormal Romberg		0.39	0.91	4.4	0.7

The role of the physical exam in diagnosing IDH has been studied extensively. Unfortunately, the gold standard for diagnosing symptomatic IDH is a complex matter and surgical findings, which have been used as the gold standard in many studies introduce selection bias. The most well-known and well-studied test for diagnosing IDH is the SLR test, although much of the literature is limited by methodological flaws.³⁴ Studies of the SLR have shown significantly heterogeneous results. This might in part be due to a lack of precision and consistency in defining a positive test. The most reliable definition appears to be reproduction of typical dermatomal pain, with a kappa of 0.68, which compares with a kappa of 0.36 for reproduction of any leg pain.²³ Pathoanatomic studies suggest that tension on the nerve root is maximal between 35 and 70° of elevation.²⁰ Thus a reasonable proposal for a standard for interpreting the SLR test as positive would be reproduction of a typical dermatomal pattern of pain with less than 70° of elevation of the passive straight leg.

Despite their heterogeneity, studies of the diagnostic utility of the SLR test have generally concluded that this test is sensitive for the presence of IDH but non-specific;

the sensitivity seems to be higher in younger patients and decrease some with age.²⁰ Sensitivities range from 0.78 to 0.97 with a pooled estimate of 0.91, and specificities from 0.10 to 0.52 with a pooled estimate of 0.26.^{13,20,34} The crossed or contralateral SLR—elicitation of typical pain in the symptomatic leg with passive raising of the straight contralateral leg—generally shows low sensitivity and higher specificity: sensitivities range from 0.22 to 0.52 with a pooled estimate of 0.29 and specificities from 0.85 to 1.0 with a pooled estimate of 0.88.^{13,20,34} The reliability of the contralateral SLR is also quite good, with a kappa of 0.7.^{23,35} The clinical utility of the SLR test can be summarized as follows: the absence of typical pain with passive SLR lowers the odds of IDH by a factor of 0.34, a positive SLR mildly increases the odds of IDH by a factor of only 1.2, the presence of a positive crossed-SLR increases the odds of IDH by a factor of 2.4. Modifications of the SLR test, including bowstring sign, foot dorsiflexion, or neck flexion, have either been unreliable or not well enough studied to assess their value.^{34,35}

The SLR tests for nerve root irritation of the lower lumbar roots, where 95% of IDH occurs, but it misses herniation in the upper lumbar spine. The femoral stretch or femoral tension sign is the equivalent test for the upper lumbar nerve roots. The femoral stretch sign has modest reliability, with a kappa of 0.3–0.5 but its diagnostic utility has not been well-studied.³⁵

Two lesser known tests for IDH have been recently studied: (i) the bell test, defined as positive when pressure applied with the thumb between L4-5 or L5-S1 spinous processes reproduced or exacerbated typical radicular pain in a standing patient; and (ii) the hyperextension test, defined as positive if typical radicular pain was reproduced with passive full lumbar extension in a standing patient with knees kept in extension.³⁶ These tests showed similar reliability to the SLR—kappas of 0.51–0.67 for the bell test, 0.35–0.5 for the hyperextension test, and 0.27–0.47 for the SLR—but their sensitivities were lower with the bell test 0.37–0.53, the hyperextension test 0.40–0.47, and the SLR 0.77–0.83. The specificity of the bell test and the hyperextension test were higher than the ipsilateral SLR—0.63 and 0.71, respectively, versus 0.37 for the SLR—but lower than the crossed-SLR, with specificity around 0.86. This study did not find a combination of tests that were more sensitive than the SLR, although the combination of a positive hyperextension test and a crossed-SLR was highly specific at 0.93 but still had an LR of only 4.0 because of limited sensitivity. Thus these findings seem to perform somewhat similarly to the SLR test but do not add much over and above the SLR.

In spinal stenosis, the physical examination appears generally less helpful than the history. The only physical examination finding with a sensitivity > 60% is the lack of pain with flexion, at 0.79, although its specificity is only 0.44; the absence of this finding only decreases the odds of SpS by half.^{9,37} Highly specific, although insensitive, findings include a wide-based gait—sensitivity 0.43 and specificity 0.97—and an abnormal Romberg—sensitivity 0.39 and specificity 0.91.^{9,37} Thigh pain with 30 seconds of lumbar extension had fairly low sensitivity and specificity (0.51 and 0.69, respectively) but was an independent predictor of a diagnosis of SpS in a regression model along with age, no pain when seated, and wide-based gait.³⁸

DOES THE PATIENT FALL INTO A SUBGROUP THAT IS MORE OR LESS LIKELY TO BENEFIT FROM SOME SPECIFIC TARGETED THERAPY?

The sacroiliac joint (SIJ), in the absence of trauma or inflammatory sacroiliitis, remains a controversial entity as a purported cause of LBP. Those skeptical of the role of

the sacroiliac joint in LBP often point to the joint's extreme strength and minimal motion.³⁹ Trials of controlled injections, however, clearly demonstrate that some individuals do have SIJ-injection-responsive pain, although the prevalence of this entity remains controversial. Some of the injection studies^{40,41} have been cited as showing that SIJ-injection-responsive pain makes up about 20% of chronic low back pain⁴², but this is extremely misleading; the finding of 20% SIJ-injection-responsive pain was in highly preselected cohorts with pain patterns felt to represent SIJ problems. The prevalence in unselected populations with chronic LBP will be lower.

The ability of the history and physical to accurately identify SIJ-injection-responsive pain remains uncertain. A systematic review has found numerous methodological shortcomings in the literature, modest reliability to most of the physical examination maneuvers, and highly variable sensitivity and specificity for each physical exam maneuver.^{43,44} The only historical feature that seems to have diagnostic utility is a lack of any pain above the level of L5; in one study only 4% of patients with SIJ-injection-responsive pain reported any pain above L5.⁴⁵ Whereas no individual physical examination maneuver has much diagnostic utility in SIJ pain, recent studies looking at clusters of tests have been somewhat more promising. A recent study looking at five SIJ provocation tests—Gaenslen's test, Patrick's (or the FABER) test, compression, distraction, and the thigh thrust test—found that a cluster of three out of the five being positive was fairly reliable, with a 94% agreement and a kappa of 0.7.⁴⁶ A similar cluster of five tests, except substituting the sacral thrust for Patrick's test, found a sensitivity of 0.91 and a specificity of between 0.78 and 0.87.⁴⁷

A recent randomized trial by Fritz et al provides some intriguing results on using history and physical examination to identify subgroups of acute LBP patients suitable for different physical therapy interventions.⁴⁸ The classification scheme does not use precise or single historic and physical exam features that can be quantified individually, but an overall classification algorithm.⁴⁹ Patients are placed into seven categories based on the treatment expected to be most beneficial: (i) immobilization (trunk strengthening exercises and possibly a brace/corset) in those patients with frequent episodes, history of frequent, transiently helpful manipulation, ligamentous laxity, improvement with brace, or 'catch' during flexion; (ii) sacral mobilization in those with asymmetric pelvis in standing, standing flexion test, positive long-sit and prone knee-bend tests; (iii) lumbar mobilization in those with unilateral lumbar pain and specific, painful limitation in active ROM; (iv) specific flexion exercises for those with preference for sitting versus standing and centralization of symptoms on repeated flexion; (v) specific extension exercises for those with preference for standing versus sitting and centralization of symptoms on repeated extension; (vi) lateral shift for those with a list and asymmetric side bending ROM; and (vii) traction in those with signs and symptoms of neural compression but who did not centralize with repeated end-range motion testing.⁴⁹ This classification was found to be moderately reliable with a 65% agreement and kappa of 0.56.⁵⁰ More importantly, when the use of this classification scheme and its specific interventions was compared with a general approach of reassurance, advice to stay active, and low-stress aerobic exercise, the classification group has significantly improved outcomes of about 11 points on the Oswestry scale and 6 points on the SF-36 PCS scale at 4 weeks, with persistent improvement of about 9 points on the Oswestry at 1 year. There was also a non-significant trend towards lower costs in the classification group.⁴⁸ These results are intriguing but come from only one relatively small study and require further investigation and confirmation.

Table 3. Diagnostic performance of imaging tests in LBP (data from Refs. 9,65,67,70).

Condition	Test	Reliability (kappa)	Sensitivity	Specificity	LR when positive	LR when negative
Malignancy	Plain radiograph		0.6	0.95–0.99	12–100	0.4
	Bone scan		0.75–0.98	0.64–0.93	4–10	0.1–0.3
	Magnetic resonance		0.83–0.93	0.9–0.97	8–31	≈ 0.1
Infective spondylitis	Plain radiograph		0.82	0.57	2	0.3
	Bone scan		0.9	0.78	4	0.1
	Magnetic resonance		0.96	0.92	12	0.04
Intervertebral disc herniation	Computed tomography	?	0.62–0.9	0.7–0.87	2–7	0.1–0.5
	Magnetic resonance	0.58	0.6–1.0	0.43–0.97	1–33	0–0.93
Spinal stenosis	Computed tomography	?	0.9	0.8–0.96	5–22	≈ 0.1
	Magnetic resonance	0.28	0.9	0.72–0.99	3–90	≈ 0.1
Instability	Plain radiograph	0.36–0.6	0.51–0.99	0.68–0.99	–	–

WHAT IS THE RELATIVE VALUE OF IMAGING IN LBP?

Modern imaging can provide important information in certain patients with low back pain (Table 3). Unfortunately, our increasing ability to image smaller and smaller details brings with it bigger and bigger problems—we can now identify a host of ‘abnormalities’ on imaging studies, which generate cascades of additional tests and treatments but might be completely unrelated to the patient’s LBP.^{1,51,52} In one study at a multidisciplinary spine center, 70% of patients had undergone advanced diagnostic imaging before referral but only 4% of these met their strict indications and only 25% met even their most lax criteria for appropriate indications.⁵³

Patients’ psychological distress has been found to be an important predictor of whether imaging, and particularly repeated imaging, is obtained, although this is unrelated to their underlying cause of LBP that the imaging is attempting to elucidate.⁵⁴ Whereas most patients with LBP think radiographic evaluation is very important⁵⁵, a randomized study showed that primary care patients receiving radiographs, while more satisfied with their care, had more doctor visits and more disability 3 months later.⁵⁶ The diagnostic value of imaging studies in LBP lies in their careful selection based on the underlying probabilities that the patient has a diagnosis that the imaging study can further elucidate.¹⁷

IS THERE A SERIOUS UNDERLYING SYSTEMIC DISEASE CAUSING THE BACK PAIN?

Imaging studies can be critical when the patient is at high risk for a serious cause of low back pain, such as tumor or infection. Standard recommendations for patients with ‘red flags’ for malignancy or infection, or who have failed to improve after 4–6 weeks of

conservative therapy, is an erythrocyte sedimentation rate (ESR) and plain radiography of the spine.¹⁴ Plain radiographs of the spine have high specificity for malignancy of 0.95–0.995 but are relatively insensitive, at 0.6.⁹ A normal radiograph has a likelihood ratio for malignancy of 0.4; however, because the starting odds are low, even with some red flags present, this might be sufficient, particularly if the ESR is also reassuring. Because radiographs have low sensitivity and MRI has high sensitivity and specificity for cancer (approximately 0.9 and 0.95, respectively)⁹, some have suggested moving directly to MRI in all patients with suspected malignancy. However, because of the low prevalence of cancer this strategy identifies <1 case per 1000 studies and the incremental cost per quality-adjusted-life-year gained has been estimated at about \$300 000.⁵⁷ A detailed cost-effectiveness analysis of different diagnostic strategies recommended imaging with MRI, or bone scan followed by MRI if the bone scan is abnormal, in patients with one or more red flags if they have either a worrisome radiograph (lytic or blastic lesion seen) or an ESR > 50; this strategy found 88% of all the findable cases at a cost on \$40 per patient, \$9525 dollars per case found, and with only 1.6 false positives per 1000 patients.⁵⁸ The one exception was patients with a personal history of cancer; because their pretest probability is 10–15%, moving directly to imaging with MRI is probably warranted.⁵⁸

Infective spondylitis can be difficult to diagnose with plain radiography. The sensitivity of plain films is about 0.82 and the specificity 0.57.⁹ Sensitivity will be even lower early on in the course. Bone scanning has good sensitivity for infection of 0.90 but modest specificity of 0.78.⁹ MRI has excellent sensitivity and specificity (0.96 and 0.92, respectively) and is the test of choice in patients with a high clinical suspicion for infective spondylitis.

Inflammation of one or both sacroiliac joints is a characteristic feature and part of the diagnostic and classification criteria for spondyloarthritis; sacroiliitis is symmetric in > 90% of AS.⁵⁹ Radiography might be normal early on and the typical duration between clinical features of inflammatory back and typical radiographic changes is in the range of 1–9 years.⁵⁹ Plain radiography is estimated to have a sensitivity of 0.25–0.45 for AS but it is highly specific.⁹ When clear evidence of erosions or frank ankylosis is present the interpretation is clear, although when more subtle changes, such as possible small erosions or sclerosis, are the only changes then interpretation becomes more uncertain.

CT scanning can detect chronic morphologic changes with greater sensitivity than plain radiographs but is not deemed useful for assessing active inflammation of the joint. Radionuclide scintigraphy has had mixed results: Some authors report it to be specific but not very sensitive⁹ and others as sensitive but very non-specific for AS.⁵⁹ Blum et al found scintigraphy to have modest sensitivity of 0.48 and high specificity of 0.97.⁶⁰ Contrast-enhanced MRI appears to be both highly sensitive and specific for active sacroiliitis (0.995 and > 0.95, respectively).^{59,60} However, the relationship between these imaging findings and clinical features remains somewhat unclear.⁶¹

IS THERE NEUROLOGICAL IMPAIRMENT THAT MIGHT REQUIRE, OR BENEFIT FROM, SURGICAL EVALUATION?

As discussed above, cauda equina syndrome is a clinical diagnosis and a surgical emergency. Imaging is not needed for the diagnosis but appropriate imaging, typically MRI, should be obtained immediately to confirm the diagnosis and plan the intervention.

Herniated discs cannot be seen on plain radiography; CT scans have moderate sensitivity of 0.6–0.9 and specificity of 0.7–0.87.⁹ MRI is the diagnostic test of choice for

evaluating for IDH because of its excellent soft tissue contrast but there are many caveats. The sensitivity of MRI for IDH has ranged in studies from 0.6 to 1.0 but is probably on the higher end of that spectrum.⁹ The problem is that MRI finds disc abnormalities in 10–60% of asymptomatic adults.⁹ MRI is accurate at finding herniated discs but cannot always indicate if the anatomic abnormality is clinically relevant. It is not critical to make the diagnosis of IDH as a cause of sciatica early because the natural history is favorable: most patients improve within the first 6 weeks and the herniated portion of the disc often resorbs over time.⁵ Thus MRI within the first 4–6 weeks of suspected IDH is necessary only if there is progressive neurological deficit or intolerable pain levels despite conservative treatment and surgical referral is being planned. The main role of MRI in this setting should be to shift patients with suspected IDH considering surgical intervention based on clinical findings towards a more conservative course if the suspected IDH and nerve root compression is not seen.⁶²

One step that probably should be taken to help improve the value of MRI is to avoid the non-specific term ‘disc herniation’ in favor of a stricter morphological classification of bulge, protrusion, extrusion, and sequestered fragment.⁶³ Disc bulges occur in over 50% of asymptomatic adults⁹ and should probably be considered normal finding. However, extrusions are rare in asymptomatic adults (1%) and can be of much greater clinical significance.^{64,65} Studies to prove this hypothesis, however, remain to be done. Some have advocated the use of gadolinium to look for nerve root enhancement but the data on the value of this approach is conflicting.⁹ MRI findings in initially asymptomatic adults do not predict the future development of LBP, or of sciatica.⁶⁶

Plain radiography is of limited value in assessing SpS. CT and MRI have shown similar sensitivity and specificity in diagnosing central SpS, both in the range of 0.9 for sensitivity and 0.8–0.95 for specificity. However, depending on the age of the cohort, anatomic SpS is quite common in asymptomatic adults (up to 28%).⁹ There is currently no standardized grading scheme for lumbar SpS and radiological interpretation can be quite variable. Speciale et al found fair to poor reliability of grading the degree of stenosis, with kappas of 0.11–0.26. Furthermore, there was large variability in the measured cross-sectional area of the spinal canal within each reader’s ratings of mild, moderate, and severe stenosis.⁶⁷ In addition, the correlation between imaging findings and clinical symptoms, both at baseline and postoperatively, is not clear.⁶⁸ As with IDH, the main role of CT or MRI should be to confirm the diagnosis in patients clinically suspected of having SpS who have failed conservative treatment and are contemplating surgery.

DOES THE PATIENT FALL INTO A SUBGROUP THAT IS MORE OR LESS LIKELY TO BENEFIT FROM SOME SPECIFIC TARGETED THERAPY?

Outside the assessment of serious underlying causes of LBP and confirming a clinical diagnosis of IDH or SpS in a patient being considered for surgical intervention, there is very little evidence to support the utility of imaging in the diagnosis or management of LBP. A systematic review of radiographic findings in LBP found that disc degeneration defined by disc space narrowing, osteophytes, and sclerosis was very weakly associated with the presence of LBP, with an odds ratio between 1.2 and 3.3; other findings, including spondylolysis, spondylolisthesis, spina bifida, transitional vertebrae, spondylolysis, and Scheurmann’s disease, were not significantly associated with the presence of LBP.⁶⁹ Shaffer et al found high false-positive and false-negative rates when assessing lumbar segmental instability by radiograph.⁷⁰

Because MRI scans give vastly improved images compared with lumbar radiographs, it is reasonable to hypothesize that replacing radiographs with MRI in LBP might be more informative. This hypothesis was tested in a randomized trial in which patients otherwise referred for lumbar radiographs were randomized to radiography or rapid MRI. Because of the frequent finding of asymptomatic abnormalities and the lack of a gold standard, the MRI group was assessed not by diagnostic accuracy per se but for whether the information gained had any impact on patient outcome.⁷¹ The MRI group had a trend towards higher rates of surgery and slightly higher costs, but without any improvement in pain or functional outcomes. MRI may be more useful in highly selected population but this needs to be tested.

Imaging has been evaluated in the diagnosis of SIJ-injection-responsive pain with consistently poor results. CT scanning had a sensitivity of 0.58 and a specificity of 0.69.⁷² Radionuclide scanning had higher specificity (in the range of 0.89–0.96) but had a consistently low sensitivity of 0.13–0.46, making it of limited clinical value.^{73,74}

Practice points

- the diagnostic approach to low back pain should assess for the possibility of serious underlying systemic disease, the presence of neurological compromise that might require or benefit from surgical referral, and, if these are absent, features that might help guide the choice of conservative therapy
- historical features are the most useful in assessing for serious underlying disease
- history and physical are the key factors in assessing for neurological compromise
- imaging, particularly MRI, serves best to assess for serious underlying disease in high-risk patients and to confirm a suspected diagnosis of intervertebral disc herniation or spinal stenosis in patients being considered for surgical interventions
- outside the setting of suspected systemic disease or neurological compromise, radiography, scintigraphy, computed-tomography, and MRI add very little to the diagnostic evaluation of low back pain

Research agenda

- further controlled trials comparing non-specific therapy for LBP to targeted therapy based on well-defined subgroups determined by history, physical examination \pm imaging are critical
- given the general lack of any gold standard for linking pathoanatomic findings with the symptoms of LBP, more studies of imaging should use the paradigm of therapeutic impact rather than traditional diagnostic accuracy
- well-designed, controlled studies assessing the clinical, psychosocial, imaging, and perhaps genetic predictors of which patients with herniated discs and spinal stenosis are most likely to benefit from surgery are needed

ACKNOWLEDGEMENT

Dr Lurie is supported by Grant K23AR48138-03 from the National Institute of Arthritis, Musculoskeletal, and Skin Diseases.

REFERENCES

1. Breslau J & Seidenwurm D. Socioeconomic aspects of spinal imaging: impact of radiological diagnosis on lumbar spine-related disability. *Topics in Magnetic Resonance Imaging* 2000; **11**(4): 218–223.
- *2. Carragee EJ & Hannibal M. Evaluation of Low Back Pain. *The Orthopedic clinics of North America* 2004; **35**: 7–16.
3. Deyo R. Cost, controversy, crisis: low back pain and the health of the public. *Annual Review of Public Health* 1991; **12**: 141–156.
4. Deyo RA. Fads in the treatment of low back pain. *The New England Journal of Medicine* 1991; **325**(14): 1039–1040.
- *5. Deyo RA & Weinstein JN. back pain. *The New England Journal of Medicine* 2001; **344**(5): 363–370.
6. Waddell G. Biopsychosocial analysis of low back pain. *Acta Orthopaedica Scandinavica. Supplementum* 1993; **251**: 21–24.
7. Dudler J & Balague F. What is the rational diagnostic approach to spinal disorders? Bilateral primary brucellar psoas abscess. *Best Practice and Research in Clinical Rheumatology* 2002; **16**(1): 43–57.
8. Cherkin DC & MacCornack F. Patient evaluations of low back pain care from family physicians and chiropractors. *Western Journal of Medicine* 1989; **150**: 351–355.
- *9. Jarvik JG & Deyo RA. Diagnostic evaluation of low back pain with emphasis on imaging. *Annals of Internal Medicine* 2002; **137**(7): 586–597.
10. Borenstein DG. A clinician's approach to acute low back pain. *American Journal of Medicine* 1997; **102**(1A): 16S–22S.
11. Lurie JD, Gerber PD & Sox HC. Clinical problem solving: a pain in the back. *New England Journal of Medicine* 2000; **343**(10): 723–726.
12. Frymoyer J. Back pain and sciatica. *New England Journal of Medicine* 1988; **318**(5): 291–300.
13. van den Hoogen HM, Koes BW, van der Heijden GJ & Bouter LM. On the accuracy of history, physical examination, and erythrocyte sedimentation rate in diagnosing low back pain in general practice. A criteria-based review of the literature. *Spine* 1995; **20**(3): 318–327 [see comment].
14. Wipf JE & Deyo RA. Low back pain. *Medical Clinics of North America* 1995; **79**(2): 231–246.
- *15. Atlas SJ & Deyo RA. Evaluating and managing acute low back pain in the primary care setting. *Journal of General Internal Medicine* 2001; **16**(2): 120–131.
16. Deyo RA. Early diagnostic evaluation of low back pain. *Journal of General Internal Medicine* 1986; **1**(5): 328–338.
- *17. Deyo J & Rainville D. What can the history and physical examination tell us about low back pain? *Journal of the American Medical Association* 1992; **268**: 760–765.
18. Sadowska-Wroblewska M, Filipowicz A, Garwolinska H, Michalski J, Rusiniak B & Wroblewska T. Clinical symptoms and signs useful in the early diagnosis of ankylosing spondylitis. *Clinical Rheumatology* 1983; **2**(1): 37–43.
19. Husby G, Hordvik M & Gran JT. An epidemiological survey of the signs and symptoms of ankylosing spondylitis. *Annals of the Rheumatic Diseases* 1985; **44**(6): 359–367.
20. Andersson GB & Deyo RA. History and physical examination in patients with herniated lumbar discs. *Spine* 1996; **21**(supplement 24): 10S–18S.
21. Vroomen PC, de Krom MC & Wilmink JT. Pathoanatomy of clinical findings in patients with sciatica: a magnetic resonance imaging study. *Journal of Neurosurgery* 2000; **92**(2): 135–141.
22. Vroomen PC, de Krom MC & Knottnerus JA. Diagnostic value of history and physical examination in patients suspected of sciatica due to disc herniation: a systematic review. *Journal of Neurology* 1999; **246**(10): 899–906.

23. Vroomen PC, de Krom MC & Knottnerus JA. Consistency of history taking and physical examination in patients with suspected lumbar nerve root involvement. *Spine* 2000; **25**(1): 91–96 [discussion 97].
24. Nitta H, Tajima T, Sugiyama H et al. Study on dermatomes by means of selective lumbar spinal nerve block. *Spine* 1993; **18**(13): 1782–1786.
25. van Tulder MW, Koes BV & Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine* 1997; **22**: 2128–2156 [see comment].
26. Bouter LM, van Tulder MW, Koes BV & van Poppel MN. Methodologic issues in low back pain research in primary care. *Spine* 1998; **23**(18): 2014–2020.
27. Borkan JM, Koes B, Reis S & Cherkin DC. An agenda for primary care research on low back pain. *Spine* 1998; **23**(18): 1992–1996.
28. Revel M, Poiraudou S, Auleley GR et al. Capacity of the clinical picture to characterize low back pain relieved by facet joint anesthesia: proposed criteria to identify patients with painful facet joints. *Spine* 1998; **23**(18): 1972–1976.
29. Schwarzer AC, Wang SC, Bogduk N et al. Prevalence and clinical features of lumbar zygapophysial joint pain: a study in an Australian population with chronic low back pain. *Annals of the Rheumatic Diseases* 1995; **54**(2): 100–106.
30. Flynn T, Fritz J, Whitman J et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine* 2002; **27**(24): 2835–2843.
- *31. Childs JD, Fritz JM, Flynn TW et al. Clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Annals of Internal Medicine* 2004; **141**: 920–928.
32. Deyo RA. Treatments for back pain: can we get past trivial effects? *Annals of Internal Medicine* 2004; **141**: 957–958.
33. Pincus T, Burton AK, Vogel S & Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 2002; **27**(5): E109–E120.
34. Deville WL, van der Windt DA, Dzaferagic A et al. The test of Lasegue: systematic review of the accuracy in diagnosing herniated discs. *Spine* 2000; **25**(9): 1140–1147.
35. McCombe PF. Volvo Award in clinical sciences. Reproducibility of physical signs in low-back pain 1989; [see comment].
36. Poiraudou S, Foltz V, Drape JL, Fermanian J, Lefevre-Colau MM, Mayoux-Benhamou MA et al. Value of the bell test and the hyperextension test for diagnosis in sciatica associated with disc herniation: comparison with Lasegue's sign and the crossed Lasegue's sign. *Rheumatology* 2001; **40**(4): 460–466.
37. Fritz JM, Delitto A, Welch WC & Erhard RE. Lumbar spinal stenosis: a review of current concepts in evaluation, management, and outcome measurements. *Archives of Physical Medicine and Rehabilitation* 1998; **79**(6): 700–708.
38. Katz JN, Dalgas M, Stucki G et al. Degenerative lumbar spinal stenosis. Diagnostic value of the history and physical examination. *Arthritis and Rheumatism* 1995; **38**(9): 1236–1241.
39. Walker JM, Wright DL & Kemp TL. The sacroiliac joint: a critical review. *Physical Therapy* 1992; **72**(12): 903–916.
40. Maigne JY, Aivaliklis A & Pfefer F. Results of sacroiliac joint double block and value of sacroiliac pain provocation tests in 54 patients with low back pain. *Spine* 1996; **21**(16): 1889–1892 [see comment].
41. Aprill CN, Derby R, Fortin J et al. The sacroiliac joint in chronic low back pain. *Spine* 1995; **20**(17): 1878–1883.
42. Bogduk N. Management of chronic low back pain. *Medical Journal of Australia* 2004; **180**(2): 79–83.
43. van der Wurff P, Meyne W & Hagmeijer RH. Clinical tests of the sacroiliac joint. A systemic methodological review. Part 2: validity. *Manual Therapy* 2000; **5**(2): 89–96.
44. van der Wurff P, Hagmeijer RH & Meyne W. Clinical tests of the sacroiliac joint. A systemic methodological review. Part 1: reliability. *Manual Therapy* 2000; **5**(1): 30–36.
45. Dreyfuss P. The value of medical history and physical examination in diagnosing sacroiliac joint pain. *Spine* 1996; **21**(22): 2594–2602 [see comment].
46. Kokmeyer DJ, Van der Wurff P, Aufdemkampe G et al. The reliability of multitest regimens with sacroiliac pain provocation tests. *Journal of Manipulative Physiological Therapeutics* 2002; **25**(1): 42–48.
47. Laslett M, Young SB, Aprill CN & McDonald B. Diagnosing painful sacroiliac joints: a validity study of a McKenzie evaluation and sacroiliac provocation tests. *Australian Journal of Physiotherapy* 2003; **49**: 89–97.

- *48. Fritz JM, Delitto A & Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine* 2003; **28**(13): 1363–1371 [discussion 1372].
49. Delitto A & Erhard RE. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Physical Therapy* 1996; **76**(2): 187–190 [see comment].
50. Fritz JM & George S. The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. *Spine* 2000; **25**(1): 106–114.
51. Deyo RA. Magnetic resonance imaging of the lumbar spine. Terrific test or tar baby? *New England Journal of Medicine* 1994; **331**(2): 115–116.
52. Lurie JD, Birkmeyer NJ & Weinstein JN. Rates of advanced spinal imaging and spine surgery. *Spine* 2003; **28**(6): 616–620.
53. Boden SD & Swanson AL. An assessment of the early management of spine problems and appropriateness of diagnostic imaging utilization. *Physical Medicine and Rehabilitation Clinics of North America* 1998; **9**(2): 411–417.
54. Selim AJ, Fincke G, Ren XS et al. Patient characteristics and patterns of use for lumbar spine radiographs: results from the Veterans Health Study. *Spine* 2000; **25**(19): 2440–2444.
55. Espeland A. Patients' views on importance and usefulness of plain radiography for low back pain. *Spine* 2001; **26**: 1356–1363 [see comment].
56. Kendrick D, Fielding K, Bentley E et al. Radiography of the lumbar spine in primary care patients with low back pain: randomised controlled trial. *British Medical Journal* 2001; **322**: 400–405 [see comment].
57. Hollingworth W, Blackmore CC, Alotis MA et al. Rapid magnetic resonance imaging for diagnosing cancer-related low back pain. *Radiology* 2003; **227**(3): 669–680.
- *58. Joines JD, McNutt RA, Carey TS et al. cancer in primary care outpatients with low back pain: a comparison of diagnostic strategies. *Journal of General Internal Medicine* 2001; **16**(1): 14–23.
59. Braun J, Rudwaleit M, Boonen A & Zink A. Imaging of sacroiliitis. *Annals of Rheumatic Diseases* 2002; **61**(supplement 3): iii8–iii18.
60. Blum U, Buitrago-Tellez C, Mundinger A et al. Magnetic Resonance Imaging for Detection of Active Sacroiliitis—A prospective study comparing conventional radiography, scintigraphy, and contrast enhanced MRI. *Journal of Rheumatology* 1996; **23**: 2107–2115.
61. Puhakka KB, Jurik AG, Schiottz-Christensen B et al. Magnetic resonance imaging of sacroiliitis in early seronegative spondylarthropathy. Abnormalities correlated to clinical and laboratory findings. *Rheumatology* 2004; **43**(2): 234–237.
62. Rankine JJ, Gill KP, Hutchinson CE et al. The therapeutic impact of lumbar spine MRI on patients with low back and leg pain. *Clinical Radiology* 1998; **53**(9): 688–693.
63. Fardon DF & Milette PC. Combined task forces of the north American Spine Society ASoS, American Society of N. Nomenclature and classification of lumbar disc pathology. Recommendations of the combined task forces of the north American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. *Spine* 2001; **26**(5): E93–E113.
64. Jarvik JJ, Hollingworth W, Heagerty P et al. The Longitudinal Assessment of Imaging and Disability of the Back (LAIDBack) Study: baseline data. *Spine* 2001; **26**(10): 1158–1166.
65. Jensen M, Brant-Zawadzki M & Obuchowski N. Magnetic resonance imaging of the lumbar spine in people without back pain. *New England Journal of Medicine* 1994; **331**: 69–73.
66. Borenstein DG, O'Mara Jr. JW, Boden SD et al. The value of magnetic resonance imaging of the lumbar spine to predict low-back pain in asymptomatic subjects: a seven-year follow-up study. *Journal of Bone and Joint Surgery—American* 2001; **83-A**(9): 1306–1311.
67. Speciale AC, Pietrobon R, Urban CW et al. Observer variability in assessing lumbar spinal stenosis severity on magnetic resonance imaging and its relation to cross-sectional spinal canal area. *Spine* 2002; **27**(10): 1082–1086.
68. Amundsen T, Weber H, Lilleas F et al. Lumbar spinal stenosis. Clinical and radiologic features. *Spine* 1995; **20**(10): 1178–1186.
- *69. 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**(4): 427–434.
70. Shaffer WO, Spratt KF, Weinstein J et al. Volvo Award in clinical sciences. The consistency and accuracy of roentgenograms for measuring sagittal translation in the lumbar vertebral motion segment. An experimental model. *Spine* 1990; **15**(8): 741–750.

71. Jarvik JJ & Jarvik JG. Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *Annals of Internal Medicine* 2003; **139**(11): 950–951 [see comment].
72. Elgafy H, Semaan HB, Ebraheim NA & Coombs RJ. Computed tomography findings in patients with sacroiliac pain. *Clinical Orthopaedics and Related Research* 2001; **382**: 112–118.
73. Maigne JY, Boulahdour H & Chatellier G. Value of quantitative radionuclide bone scanning in the diagnosis of sacroiliac joint syndrome in 32 patients with low back pain. *European Spine Journal* 1998; **7**: 328–331.
74. Slipman CW, Sterenfeld EB, Chou LH et al. The value of radionuclide imaging in the diagnosis of sacroiliac joint syndrome. *Spine* 1996; **21**(19): 2251–2254.