

Screening for malignancy in low back pain patients: a systematic review

Nicholas Henschke · Christopher G. Maher · Kathryn M. Refshauge

Received: 19 March 2007 / Accepted: 21 May 2007 / Published online: 14 June 2007
© Springer-Verlag 2007

Abstract To describe the accuracy of clinical features and tests used to screen for malignancy in patients with low back pain. A systematic review was performed on all available records on MEDLINE, EMBASE, and CINAHL electronic databases. Studies were considered eligible if they investigated a cohort of low back pain patients, used an appropriate reference standard, and reported sufficient data on the diagnostic accuracy of tests. Two authors independently assessed methodological quality and extracted data to calculate positive (LR+) and negative (LR–) likelihood ratios. Six studies evaluating 22 different clinical features and tests were identified. The prevalence of malignancy ranged from 0.1 to 3.5%. A previous history of cancer (LR+ = 23.7), elevated ESR (LR+ = 18.0), reduced hematocrit (LR+ = 18.2), and overall clinician judgement (LR+ = 12.1) increased the probability of malignancy when present. A combination of age ≥ 50 years, a previous history of cancer, unexplained weight loss, and failure to improve after 1 month had a reported sensitivity of 100%. Overall, there was poor reporting of methodological quality items, and very few studies were performed in community primary care settings. Malignancy is rare as a cause of low back pain. The most useful features and tests are a previous history of cancer, elevated ESR, reduced hematocrit, and clinician judgement.

Keywords Low back pain · Diagnosis · Malignancy · Red flags

Introduction

Low back pain is one of the most common complaints in primary care. The great majority of low back pain is benign in nature, and specific diagnoses are rarely made [17]. The main purpose of the primary care assessment is to identify those cases where low back pain is caused by serious spinal pathology such as vertebral fracture, malignancy, infection, or inflammatory disease [13].

Whilst malignancy is the most common of these serious diseases, it is estimated to occur in less than 1% of primary care patients with low back pain [7]. However, early detection and treatment of spinal malignancies is important to prevent the spread of any metastatic disease and the development of further complications such as spinal cord compression [15]. The consequences of late or missed diagnosis of spinal malignancy necessitates use of accurate screening tools in primary care. Ideally, primary care practitioners should be able to identify the small number of patients with spinal malignancy without subjecting a large proportion of their low back pain patients to unnecessary diagnostic testing [11].

Clinical guidelines for the management of low back pain recommend the use of “red flag” screening questions to alert clinicians to the presence of serious disease, and indicate when further investigation is required [13]. The evidence for using these “red flags” is often based on single studies [1] or simply referenced to previous guidelines in which there was no evidence [17]. Most clinical features considered to be “red flags” for malignancy in low back pain are derived from the study performed by Deyo and Diehl in 1988 [6].

Because of the importance of identifying patients with low back pain caused by spinal malignancy in primary care and the relative lack of data in the clinical guidelines, we

N. Henschke (✉) · C. G. Maher · K. M. Refshauge
Back Pain Research Group, School of Physiotherapy,
University of Sydney, PO Box 170, Lidcombe,
NSW 1825, Australia
e-mail: N.Henschke@fhs.usyd.edu.au

performed a systematic review. We aimed to describe the diagnostic accuracy of tests used in primary care to screen for spinal malignancy in patients with low back pain.

Materials and methods

Data sources

A comprehensive search of the literature was performed to identify all relevant original, peer-reviewed articles evaluating tests for spinal malignancy in patients presenting with low back pain. The primary search was performed from the earliest available dates to 15th August 2006, on the MEDLINE, EMBASE, and CINAHL electronic databases. A subject-specific search strategy was used, combining sensitive searches of the diagnostic (index) tests available to primary care practitioners, and the target disease (low back pain) [4] (Appendix 1). The index tests included information from the history and physical examination, diagnostic imaging, and laboratory tests. Non-English language reports were included, but articles were excluded from analysis if appropriate translation was not available.

From the results of the electronic search, the bibliographies of all systematic reviews and eligible diagnostic and screening studies were reviewed. Eligible studies were entered into Web of Science to identify any articles in which they had been cited. Contact was made with experts on diagnostic testing, and on low back pain, to identify unpublished studies missed by the search process and to review the list of identified studies to ensure the search was comprehensive.

Study selection

The titles of the studies identified by the search were screened in order to exclude those that were clearly outside the scope of the review. To determine eligibility for the analysis, studies were included if they satisfied the following criteria: (a) reported on a cohort of patients presenting for either treatment for low back pain or lumbar spine X-rays; (b) confirmed the diagnosis of malignancy with an appropriate reference standard; (c) evaluated the diagnostic performance of a test available to primary care practitioners; and (d) reported results in sufficient detail to allow reconstruction of contingency tables of the raw data.

Study quality assessment

There are several potential threats to internal and external validity in studies of diagnostic accuracy [2]. Studies with

methodological shortcomings may overestimate the accuracy of a diagnostic test [3] therefore, all eligible studies identified by the search underwent methodological quality assessment using the QUADAS scale [18].

Data extraction

Two authors independently extracted the following data from each eligible article; author(s), year, journal, setting (i.e., primary care, secondary care), index tests, reference standard, number of patients, prevalence of cancer, true-positive, true-negative, false-positive, and false-negative results for the index tests. Disagreements were resolved via discussion and consensus. Because there were empty cells in the contingency table, a value of 0.5 was added to each cell in order to circumvent computational problems [10]. From the extracted data, sensitivity, specificity, and positive (LR+) and negative (LR-) likelihood ratios with their 95% confidence intervals (95%CI) were calculated using Meta-DiSc software [19]. We considered clinical features to be useful for raising the index of suspicion of malignancy if the LR+ and lower bound of the 95%CI were greater than 1. Conversely, the LR- were considered useful to lower the suspicion if the point estimate and upper bound of the 95%CI were below 1. It was our intention to pool the results and perform a meta-analysis if sufficient statistical and clinical homogeneity existed amongst the studies. If insufficient data were reported in the articles, we contacted the authors of the original studies in order to gain access to the primary data.

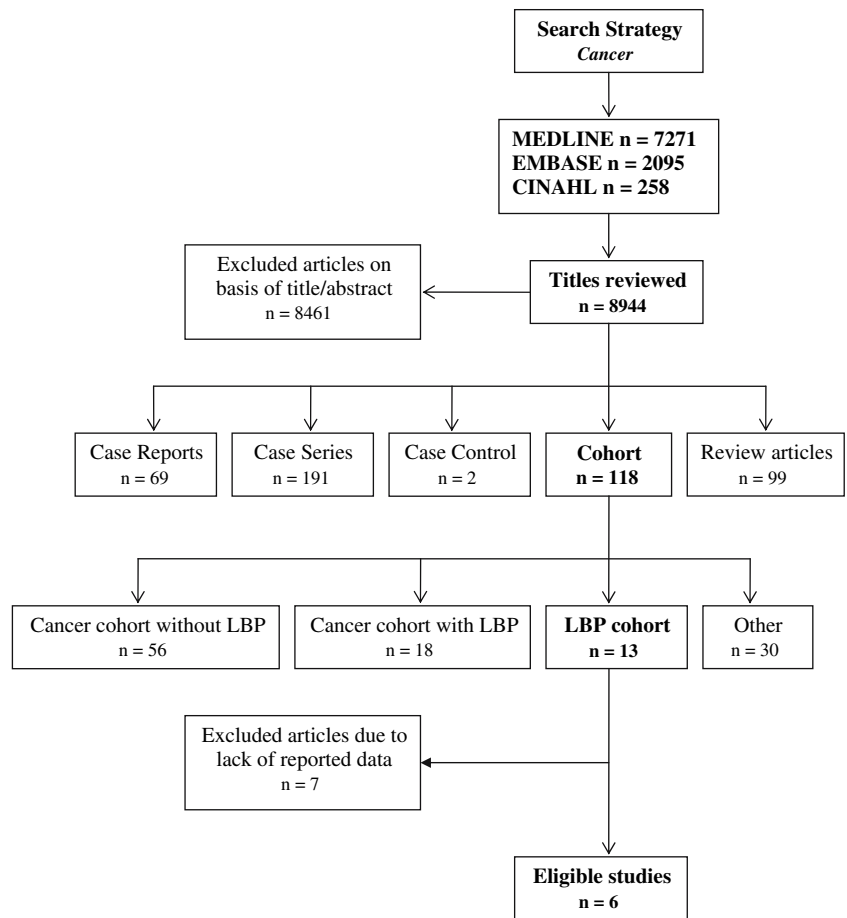
Results

Search results

The search of the electronic databases retrieved 8,944 articles (Fig. 1). After review of the titles, 8,461 articles were excluded because they were clearly ineligible. The remaining studies were categorised according to study type to screen out any reviews, case series, case reports, and case-control studies. Two authors reviewed the titles and abstracts of the cohort studies to identify all studies evaluating a cohort of low back pain patients. Any discrepancies were resolved by reading the full text and subsequent consensus. Four systematic reviews were identified by the search, and were read to identify any eligible studies missed by the search strategy.

The full text of the 13 studies that investigated cohorts of low back pain patients were read by two authors, and assessed for eligibility. Only six studies assessed tests available to primary care clinicians for the diagnosis of

Fig. 1 Study selection process and reasons for exclusion



malignancy and reported data in sufficient detail for analysis [5, 6, 8, 9, 12, 16].

Study characteristics

The six eligible studies assessed a total of 5,097 patients presenting for low back pain treatment or lumbar spine

X-rays (Table 1). The prevalence of malignancy in these studies ranged from 0.1 [12] to 3.5% [8]. Three of the eligible studies recruited patients seeking low back pain treatment from walk-in hospital clinics [5, 6, 9]. The other studies reported on patients recruited from secondary referral centres [12], patients presenting to an accident and emergency department [16], or from the office of an

Table 1 Study characteristics

| Authors | Design | Setting | Patients | Prevalence of cancer (%) (n) | Reference standard |
|----------------------------|----------------------------|---|--|------------------------------|--------------------------|
| Deyo and Diehl (1988) [6] | Prospective | Walk in clinic of public hospital | 1,975 patients seeking treatment for low back pain | 0.7 (13) | X-ray, follow-up |
| Deyo and Diehl (1986) [5] | Prospective | Walk-in clinic of university hospital | 621 patients seeking treatment for LBP | 0.6 (4) | X-ray, follow-up |
| Fernbach et al. (1976) [8] | Retrospective chart review | Practice of an orthopaedic surgeon | 518 patients (259 aged ≥50 years, 259 aged <50 years) with low back pain | 3.5 (18) | Final clinical diagnosis |
| Frazier et al. (1989) [9] | Retrospective chart review | Walk-in clinics of three teaching hospitals | 471 patients with acute low back pain | 0.4 (2) | Final clinical diagnosis |
| Khoo et al. (2003) [12] | Prospective | Hospital radiology department | 1,030 general practice referrals for lumbar spine X-rays | 0.1 (1) | X-ray |
| Reinus et al. (1998) [16] | Prospective, consecutive | Accident and emergency department | 482 patients receiving lumbar spine X-rays | 1.5 (7) | X-ray |

Table 2 Study quality assessment using QUADAS scale

| Authors | Spectrum composition | Selection criteria | Disease progression bias | Partial verification bias | Differential verification bias | Incorporation bias | Test execution details | Reference execution details | Test review bias | Diagnostic review bias | Clinical review bias | Uninterpretable results | Withdrawals |
|---------------------------|----------------------|--------------------|--------------------------|---------------------------|--------------------------------|--------------------|------------------------|-----------------------------|------------------|------------------------|----------------------|-------------------------|-------------|
| Deyo and Diehl (1988) [5] | + | ? | ? | ? | ? | + | + | + | + | ? | + | - | - |
| Deyo and Diehl (1986) [5] | + | - | + | - | ? | + | + | - | ? | ? | + | - | ? |
| Fembach et al. (1976) [8] | - | + | ? | ? | - | + | + | - | ? | ? | + | - | - |
| Frazier et al. (1989) [9] | + | ? | - | - | - | + | + | - | ? | ? | ? | - | - |
| Khoo et al. (2003) [12] | ? | ? | ? | + | + | ? | - | - | ? | ? | + | - | - |
| Reinus et al. (1998) [16] | + | + | + | ? | ? | + | - | - | + | + | ? | - | - |

+ Yes, - no, ? unclear

orthopaedic surgeon [8]. The most common reference standard used in the studies was X-ray, although the retrospective studies used the final clinical diagnosis as the reference standard [8, 9]. Two studies also used a 6-month follow-up to identify patients with malignancy who may not have received an X-ray [5, 6].

Study quality assessment

To be eligible for this review, studies needed to have used an appropriate reference standard; hence this item was not included in the quality assessment table (Table 2). Most studies were either of poor quality or poorly reported, fulfilling between two and six of the 13 criteria. Inadequate reporting was a problem in all of the studies, with no study reporting sufficient information to determine if all criteria had been met. There was poor reporting of the details of the index tests and the reference standard, and whether the tests were interpreted in a blinded fashion. Most studies were subject to partial verification bias, as they failed to perform the same reference standard on the entire cohort or on a random sample of patients.

Index test results

Data on a total of 22 different clinical features were extracted from the 6 eligible studies (Table 3). Four features were investigated by more than one study; age >50 years, a previous history of cancer, not improved after 1 month, and clinician judgement. The results for these features were pooled and are also presented in Table 3.

The features investigated can be separated into features from the clinical assessment (both history and physical examination) of the patient, or results of laboratory testing. For the history and physical examination features, only age ≥ 50 years (LR- = 0.34) [5, 6, 8, 9] had a significant LR-. A number of features had significant LR+, including a previous history of cancer (pooled estimate from two studies = 23.7); failure to improve after 1 month (pooled estimate from two studies: LR+ = 3.0), no relief with bed rest (LR+ = 1.7), and duration of pain >1 month (LR+ = 2.6) [5, 6, 8, 9] age ≥ 50 years (pooled estimate from four studies: LR+ = 2.2).

The use of some laboratory-based test results had significant likelihood ratios, such as erythrocyte sedimentation rate (ESR) ≥ 50 mm/h (LR+ = 18.0; LR- = 0.46), the presence of anaemia (LR+ = 3.9; LR- = 0.53), haematocrit < 30% (LR+ = 18.2), and white blood cell count (WBC) $\geq 12,000$ (LR+ = 4.1) [6].

The accuracy of clinician judgement in the identification of patients with malignancy was assessed by two studies and had LR+ (95%CI) of 11.9 (4.8–29.6) in an accident and

Table 3 Clinical features and data extracted from eligible studies

| Clinical features | LR+ (95% CI) | LR- (95% CI) | Sensitivity (%) | Specificity (%) |
|---|-------------------|------------------|-----------------|-----------------|
| Age \geq 50 [6] | 2.6 (1.9–3.5) | 0.35 (0.14–0.87) | 77 | 71 |
| Age > 50 [5] | 2.3 (1.3–4.2) | 0.43 (0.11–1.64) | 75 | 70 |
| Age \geq 50 [8] | 1.9 (1.6–2.2) | 0.15 (0.03–0.71) | 94 | 52 |
| Age \geq 50 [9] | 1.9 (0.6–6.0) | 0.68 (0.22–2.10) | 50 | 74 |
| Age \geq 50 (pooled estimate) | 2.2 (1.8–2.7) | 0.34 (0.17–0.68) | 84 | 69 |
| Previous history of cancer [6] | 15.8 (6.9–35.8) | 0.69 (0.48–0.99) | 31 | 98 |
| Previous history of cancer [16] | 31.7 (18.2–55.2) | 0.06 (0.00–0.94) | 100 | 97 |
| Previous history of cancer (pooled estimate) | 23.7 (11.3–49.4) | 0.25 (0.01–9.19) | 55 | 98 |
| Not improved after 1 month [6] | 3.2 (1.5–7.0) | 0.75 (0.53–1.08) | 31 | 90 |
| Not improved after 1 month [5] | 3.1 (0.8–12.1) | 0.78 (0.44–1.38) | 25 | 90 |
| Not improved after 1 month (pooled estimate) | 3.0 (1.4–6.3) | 0.79 (0.58–1.07) | 29 | 90 |
| Unexplained weight loss [6] | 3.0 (1.0–9.3) | 0.87 (0.68–1.12) | 15 | 94 |
| No relief with bed rest [6] | 1.7 (1.2–2.2) | 0.22 (0.02–3.02) | 100 | 46 |
| Insidious onset [6] | 1.0 (0.7–1.6) | 0.94 (0.49–1.80) | 62 | 42 |
| Duration >1 month [6] | 2.6 (1.5–4.6) | 0.62 (0.36–1.06) | 50 | 81 |
| Recent back injury [6] | 0.2 (0.0–3.0) | 1.18 (1.06–1.30) | 0 | 82 |
| Thoracic pain [6] | 1.2 (0.4–3.7) | 0.96 (0.74–1.26) | 17 | 84 |
| Severe pain [6] | 1.7 (0.7–4.2) | 0.88 (0.65–1.20) | 23 | 85 |
| Muscle spasm [6] | 0.5 (0.2–1.6) | 1.25 (0.97–1.59) | 15 | 66 |
| Spine tenderness [6] | 0.4 (0.1–1.4) | 1.37 (1.07–1.75) | 15 | 60 |
| Neuromotor deficit [6] | 0.4 (0.0–6.5) | 1.06 (0.95–1.18) | 0 | 91 |
| Fever (temp \geq 100 F) [6] | 1.8 (0.1–27.2) | 0.98 (0.89–1.09) | 0 | 98 |
| Neurological symptoms [12] | 7.5 (0.7–84.2) | 0.78 (0.35–1.73) | 0 | 97 |
| ESR \geq 20 mm/h [6] | 2.3 (1.6–3.3) | 0.37 (0.13–1.09) | 78 | 67 |
| ESR \geq 50 mm/h [6] | 18.0 (9.3–34.9) | 0.46 (0.23–0.92) | 56 | 97 |
| ESR \geq 100 mm/h [6] | 55.6 (13.5–228.7) | 0.75 (0.53–1.08) | 22 | 100 |
| Anaemia [6] | 3.9 (2.3–6.7) | 0.53 (0.29–0.99) | 55 | 86 |
| Hematocrit < 30% [6] | 18.2 (3.5–95.4) | 0.88 (0.71–1.09) | 9 | 99 |
| WBC \geq 12,000 [6] | 4.1 (1.3–12.5) | 0.80 (0.56–1.14) | 22 | 94 |
| Clinician judgement [12] | 12.6 (1.1–143.9) | 0.77 (0.34–1.70) | 0 | 98 |
| Clinician judgement [16] | 11.9 (4.8–29.6) | 0.58 (0.32–1.08) | 43 | 96 |
| Clinician judgement (pooled estimate) | 12.1 (4.9–29.8) | 0.66 (0.40–1.08) | 38 | 98 |
| Age \geq 50 or unexplained weight loss or previous history of cancer or failure to improve over 1 month [6] | 2.4 (2.1–2.7) | 0.06 (0.00–0.91) | 100 | 60 |

Values in italic represent significant likelihood ratios and 95% confidence intervals

LR likelihood ratio, ESR erythrocyte sedimentation rate, WBC white blood cell count

emergency setting [16], and 12.6 (1.1–143.9) in a secondary referral centre [12].

One study reported on a combination of features; age >50 years or unexplained weight loss or a past history of cancer or no improvement in low back pain after a month. This combination had a reported sensitivity of 100% [6], and a specificity of 60%, which was reported in a subsequent paper [7]. The LR+ (95%CI) was 2.4 (2.1–2.7) and the LR- (95%CI) was 0.06 (0.00–0.91).

Discussion

Using clinical features or tests to screen for serious pathologies in low back pain patients involves identifying features which, when present, raise the index of suspicion and when absent, lower the index of suspicion of having the disease. For malignancy in particular, raising the index of suspicion is most important due to the prevalence of the disease within this patient group being around 1%. The

Table 4 Application of red flags to clinical decision making

| | Pre-test probability 1% | Pre-test probability 5% |
|---|---------------------------------|-------------------------------|
| | Post-test probability (%) | |
| Clinical feature | | |
| No relief with bed rest | 1.7 | 8.3 |
| Age ≥ 50 | 2.2 | 10.4 |
| Duration of pain >1 month | 2.5 | 12.1 |
| Not improved after 1 month | 2.9 | 13.7 |
| Previous history of cancer | 19.2 | 55.7 |
| Laboratory test result | | |
| Anaemia | 3.8 | 17.1 |
| WBC $\geq 12,000$ | 3.9 | 17.9 |
| Hematocrit $< 30\%$ | 15.4 | 49.1 |
| ESR ≥ 20 mm/h | 2.3 | 10.9 |
| ESR ≥ 50 mm/h | 15.3 | 48.8 |
| ESR ≥ 100 mm/h | 35.7 | 74.7 |
| Positive clinician judgement | 10.8 | 39.1 |
| Age ≥ 50 or unexplained weight loss or previous history of cancer or failure to improve over 1 month | 2.3 | 11.3 |

Changes in the probability of cancer with a positive response to each red flag. Analysis is conducted for pre-test probabilities of 1 and 5% ESR erythrocyte sedimentation rate, WBC white blood cell count

results of this systematic review identified a number of features, which raise the probability of malignancy, however these features are not equally useful for this purpose. The LR+ of the features ranged from 1.7 to 55.6 and this needs to be appreciated when judging the clinical importance of a red flag identified in a clinical assessment.

Age ≥ 50 years, no improvement after 1 month, a previous history of cancer, and no relief with bed rest are commonly suggested “red flags” for malignancy in clinical guidelines [17], and are supported by the results of this review. Of these four red flags, a previous history of cancer is the most informative with a pooled LR+ of 23.7. The other three all had LR+ about 3. Other common “red flags” include unexplained weight loss, fever, thoracic pain, or being systematically unwell [17]. Being systematically unwell was not evaluated by any of the eligible studies, and the other features did not significantly raise or lower the probability of having malignancy [6].

While laboratory tests are not recommended routinely in low back pain patients [13, 17] tests for ESR and anaemia were found to be useful screening tools for malignancy. Hematocrit $<30\%$ (LR+ = 18.2) and WBC $\geq 12,000$ (LR+ = 4.2) also significantly raise the suspicion of malignancy [6]. In the study, which evaluated these

laboratory tests, however, the decision to perform them was based on clinician judgement [6] and the results would therefore be subject to a form of filter bias [14]. Overall clinician judgement for the presence of malignancy also had significant LR+ of 12.1 [16] but the details of what factors and other features were contained within this overall judgement were not reported.

Providing data, such as likelihood ratios, on the diagnostic accuracy of clinical features to screen for malignancy allows clinicians to evaluate whether further testing is warranted in patients with low back pain. The results of this review show that whilst a number of features have significant likelihood ratios, only four features; a previous history of cancer, an elevated ESR, low hematocrit, and clinician judgement are able to raise the post-test probability of malignancy to a clinically significant level when used in isolation (Table 4). This process is illustrated in Table 4, which shows the post-test probability of cancer in patients with a positive response to each red flag. The analysis is conducted for pre-test probabilities of 1 and 5%. For example, if the prevalence of malignancy (pre-test probability) in a low back pain patient is presumed to be 1%, and the patient is aged ≥ 50 years, the (post-test) probability would only increase to 2.2%. In fact all but one of the red flags from the clinical assessment had only modest predictive ability. The exception is if a patient has a previous history of cancer, where the probability will be raised to 19.2%, a change in disease probability that would be sufficiently large to warrant further investigation.

Clearly it would be helpful to have a clinical screening tool with greater accuracy than the clinical red flags in Table 4. One strategy would be to rely upon combinations of red flags an approach more analogous to overall clinician judgement. The only combination of features that was evaluated had a significant LR+ of only 2.4 and a significant LR- of 0.06, as the focus was on increasing the sensitivity [6]. Further study is needed which focuses on raising the suspicion of malignancy by investigating to what effect combinations of features can increase the post-test probability. Almost three-quarters of the clinical features identified by this review were investigated in only one study [6], and it is possible that other features not previously evaluated may be useful in the diagnosis of malignancy. Due to the low prevalence of the disease, large-scale high quality studies need to be performed for practitioners to have further confidence in their ability to screen for serious pathologies such as malignancy. Another area of research would be the investigation of the salient features that are considered when clinicians form an overall judgement that the patient may have cancer especially as this ‘test’ was the second most informative clinical test to identify patients with cancer. This test was found to be quite informative in two studies but neither outlined the

cues the clinicians were considering when forming this judgement.

The quality of the studies included in the review is an important consideration because certain methodological shortcomings can have large effects on estimates of diagnostic accuracy [14]. The largest of these effects are caused by studying a non-representative sample of patients, or failing to apply the same reference standard to the entire cohort or a random sample of the population [3]. Only one eligible study reported performing the same reference standard (X-ray) on all patients in their cohort [12]. The other studies combined the use of X-ray as a reference standard with clinical follow-up [5, 6, 8, 9, 16]. As clinical follow-up may fail to identify false-negative test results, the diagnostic performance of the test will be overestimated [14]. Overall, the reporting of design-related characteristics of the studies was poor, and the methodological quality was low.

To increase the external validity of our findings, we excluded case-control studies, and only extracted data from studies of clinical populations of low back pain patients. The use of clinical features for detecting serious spinal pathology is presumably most useful in the community primary care setting as this is where patients with low back pain are usually managed [13]. However, there were no studies identified by this review that were performed on a consecutive series of low back pain patients presenting to community primary care providers.

In conclusion malignancy is rare in low back patients. The most informative tests to screen for malignancy are a previous history of cancer, overall clinician judgement, elevated ESR, and reduced hematocrit. Popular red flags such as unexplained weight loss, age >50, and failure to improve after 1 month have only modest predictive ability and on their own are not useful to screen for cancer.

Acknowledgments Nicholas Henschke is under scholarship awarded by the National Health and Medical Research Council of Australia. Christopher G. Maher is a senior research fellow funded by the National Health and Medical Research Council of Australia. Nicholas Henschke will act as guarantor for the paper.

References

- Bigos S, Bowyer O, Braen G (1994) Acute low back problems in adults. Clinical practice guideline no. 14. Agency for Health Care Policy and Research, Public Health Service, US Department of Health and Human Services, Rockville, MD
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG (2003) The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 138:W1–W12
- de Vet HC, van der Weijden T, Muris JW, Heyrman J, Buntinx F, Knottnerus JA (2001) Systematic reviews of diagnostic research. Considerations about assessment and incorporation of methodological quality. *Eur J Epidemiol* 17:301–306
- Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, Bezemer PD (2002) Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2(1):92–99
- Deyo RA, Diehl AK (1986) Lumbar spine films in primary care: current use and effects of selective ordering criteria. *J Gen Intern Med* 1:20–25
- Deyo RA, Diehl AK (1988) Cancer as a cause of back pain: frequency, clinical presentation, and diagnostic strategies. *J Gen Intern Med* 3:230–238
- Deyo RA, Rainville J, Kent DL (1992) What can the history and physical examination tell us about low back pain? *JAMA* 268:760–765
- Fernbach JC, Langer F, Gross AE (1976) The significance of low back pain in older adults. *Can Med Assoc J* 115:898–900
- Frazier LM, Carey TS, Lyles MF, Khayrallah MA, McGaghie WC (1989) Selective criteria may increase lumbosacral spine roentgenogram use in acute low-back pain. *Arch Intern Med* 149:47–50
- Irwig L, Macaskill P, Glasziou P, Fahey M (1995) Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 48:119–30 (discussion 31–2)
- Joines JD, McNutt RA, Carey TS, Deyo RA, Rouhani R (2001) Finding cancer in primary care outpatients with low back pain: a comparison of diagnostic strategies. *J Gen Intern Med* 16:14–23
- Khoo LA, Heron C, Patel U, Given-Wilson R, Grundy A, Khaw KT, Dundas D (2003) The diagnostic contribution of the frontal lumbar spine radiograph in community referred low back pain—a prospective study of 1,030 patients. *Clin Radiol* 58:606–609
- Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G (2001) Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine* 26:2504–2513
- Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, Bossuyt PM (1999) Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 282:1061–1066
- Loblaw DA, Perry J, Chambers A, Laperriere NJ (2005) Systematic review of the diagnosis and management of malignant extradural spinal cord compression: the Cancer Care Ontario Practice Guidelines Initiative's Neuro-Oncology Disease Site Group. *J Clin Oncol* 23:2028–2037
- Reinus WR, Strome G, Zwemer FL Jr (1998) Use of lumbosacral spine radiographs in a level II emergency department. *Am J Roent* 170:443–447
- van Tulder M, Becker A, Bekkering T, Breen A, Gil del Real MT, Hutchinson A, Koes B, Laerum E, Malmivaara A (2004) European guidelines for the management of acute nonspecific low back pain in primary care. European Commission, Geneva. Available at <http://www.backpaineurope.org>, accessed 1st May, 2005
- Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 3:25
- Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A (2006) Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 6:31