

***Most red flags for malignancy in low back pain guidelines lack empirical support: a systematic review***

***Authors:***

Arianne P Verhagen (1), Aron Downie (2,3), Chris G Maher (2), Bart W Koes (1).

***Affiliation:***

- 1) Department of General Practice, Erasmus Medical Centre University, Rotterdam, The Netherlands
- 2) The George Institute for Global Health, Sydney Medical School, The University of Sydney
- 3) Faculty of Science and Engineering, Macquarie University, Sydney, Australia

**Word count: abstract: 198; text: 4807**

***Contact address:***

Arianne P Verhagen, Dept of General Practice, ErasmusMC, PO box 2040, 3000 CA Rotterdam, The Netherlands; tel: +3110 7044109; email: a.verhagen@erasmusmc.nl

## **Abstract**

Clinicians do not want to miss underlying serious pathology, but it is still unclear which red flags are relevant. We aimed to evaluate the origin and evidence on diagnostic accuracy of red flags for malignancy for management of low back pain (LBP) in primary care. We performed a comprehensive overview and searched the literature using snowballing techniques and reference checking for evidence on red flags endorsed in clinical guidelines for identifying patients with higher likelihood of malignancy. We selected studies including people with LBP without any restriction on study design. We extracted data on prevalence and diagnostic accuracy. Furthermore, we assessed the methodological quality of studies evaluating diagnostic accuracy. We identified 13 red flags endorsed in a total of 16 guidelines and 2 extra red flags not endorsed in any guideline. We included 33 publications varying from systematic reviews to case reports. The origin of many red flags was unclear or was sourced from case reports. The incidence of malignancy in patients presenting with LBP in primary care varied between 0% to 0.7%. Seven studies provided diagnostic accuracy data on red flags. We found 5 red flags with accuracy data from 2 or more studies, with 2 (“history of malignancy” and “strong clinical suspicion”) considered informative. In conclusion, the origin and diagnostic accuracy of many red flags endorsed in guidelines are unclear. A “history of malignancy” and “strong clinical suspicion” are the only red flags with empirical evidence of acceptable high diagnostic accuracy.

## **Key words**

Low Back Pain, Practice Guidelines/ Clinical Guidelines, red flags, diagnostic accuracy

## 1. Introduction

Low back pain (LBP) remains a common condition with an estimated lifetime prevalence of about 80% [1,2]. In 1 to 5% of patients presenting with LBP to primary care, a serious underlying pathology might be the cause, which needs timely and accurate diagnosis [4]. Spinal malignancy is, however, rare with less than 1% of patients presenting to primary care with LBP found to have a malignancy (primary vertebral tumor or vertebral metastasis) [5,6]

Clinicians are encouraged by guidelines for managing patients with LBP to screen for red flags that might indicate underlying serious pathology [7]. Most guidelines suggest that the presence of red flags justifies further investigation, mainly magnetic resonance imaging (MRI), but the threshold for testing is often unclear [8]. The Cochrane review “Red flags to screen for malignancy in patients with low back pain” evaluated diagnostic accuracy of suggested red flags and found that evidence for the use and recommendation of most red flags is weak. Without rigorous evidence on the diagnostic accuracy of red flags, this advice may be associated with unwanted side effects through unnecessary imaging (increase of radiation and health care costs), unnecessarily alarming patients (resulting in reduction of quality of life) and/or unnecessary treatment (including unnecessary surgery) [9].

A previous study demonstrated that 16 clinical guidelines for LBP, which recommended a total of 14 different red flags (including elevated ESR (erythrocyte sedimentation rate) related to the identification of malignancy [10]. Most guidelines did not justify their choice of red flags with supporting data. Therefore, the purpose of this study was to present data on the prevalence of malignancy in patients presenting with LBP and to identify the (1) origin and rationale for endorsed red flags, (2) quality of research that has evaluated red flags, and (3) diagnostic accuracy of red flags endorsed in the various guidelines for the diagnosis of malignancy in patients presenting with LBP.

## 2. Method

2.1. Design: This article was to perform a systematic review of studies evaluating the origin and diagnostic accuracy of guideline-endorsed red flags for malignancy in patients with LBP.

2.2. Search strategy: Two authors (APV, AD) independently searched for studies concerning people with LBP. Eligible studies included systematic and narrative reviews, diagnostic accuracy studies, cohorts, case control studies, and case-series considering red flags for malignancies. We searched the electronic databases (Medline and Cinahl using MESH terms (key words: low back pain, red flags, serious pathology)) from inception to on July 27, 2016. We also searched for references in the guidelines and

relevant articles found, and we performed reference checking and “snowballing” of landmark articles [11].

2.3. Selection of red flags. We selected the 13 distinct red flags for malignancies (excluding elevated ERS), which were endorsed in the 16 guidelines for LBP [10]. We defined red flags as signs or symptoms collected in the clinical assessment signalling underlying serious pathology that requires attention (Merriam-Webster dictionary).

2.4. Methodological quality. Two authors (AD, APV) independently assessed the methodological quality of all included studies that presented sensitivity or specificity data on the diagnostic accuracy of red flags using QUADAS-2 [12]. The QUADAS-2 consists of 4 key domains covering patient selection, index test, reference test and flow of patients through the study and timing of the index test(s) and reference standard. We assessed each domain in terms the risk of bias (RoB) and the first 3 domains also in terms of concerns regarding applicability. When studies scored low risk of bias on at least 2 RoB domains and unclear on the others, we considered these studies to have low RoB. When studies scored high on at least 1 RoB domain or all domains as unclear, we considered these studies to have moderate RoB. We considered that all other studies were to have high RoB.

2.5. Data extraction. Two authors (AD, APV) performed data extraction. Discrepancies were discussed until consensus was reached.

2.6. Analysis. We described prevalences of spinal malignancies in patients presenting with LBP and presented the 13 red flags in a table with findings about their origin and research evidence. In case we found additional red flags with data about diagnostic accuracy and the original authors regard this red flag as relevant (ie, the authors of the original studies stated that this red flag needs to be considered in patient care), we added them to the table.

Next, we calculate sensitivities, specificities, and likelihood ratios (LRs) of the studies that provided data. We graded the overall quality of the evidence according to the guidelines of the GRADE working groups. The quality of the evidence is rated as “high” and downgraded when one or more of the following criteria are not met: design (>50% of the study population in studies of high or moderate RoB); indirectness (>50% of the study population not originating from primary care); inconsistency (unexplained inconsistency of sensitivity and specificity in 2 or more studies); impreciseness (wide confidence intervals [Cis]); or other biases such as publication bias or fatal flaw. We did not grade the evidence when only 1 study presented data. We identified potential clinically informative red flags by a positive LR (LR+) of >15, which considers both the seriousness of the disease and very low pretest probability.

### **3. Results**

3.2. Search results. The search resulted in 164 articles and checking the reference list of the 16 guidelines and snowballing the landmark paper of Deyo & Diehl (1988) resulted in 26 references. Full text screening was done on 63 references, which were judged relevant based on title and abstract. Finally, we included 33 articles, see flow chart (figure 1). Of the 26 references found using snowballing, 6 studies, all case reports, were judged relevant.

We found two systematic reviews presented in 5 articles [13-17]. Furthermore, we found 2 narrative reviews [10,18], 7 prospective cohort studies [19-26], 9 retrospective cohort studies often collecting data from medical records [27-35], and 8 case reports [36-43]. Six studies provided diagnostic test accuracy data related to red flags for detection of malignancy for patients with LBP and were included in the assessment of diagnostic accuracy using the QUADAS-2 (Table 1)[20,21,22,24,26,29].

#### 3.2. Prevalence

The prevalence of spinal malignancies in patients with LBP seemed to be influenced by setting. The prospective primary care cohort studies (4 studies; 4864 participants) reported prevalences of malignancy in patients with LBP between 0% to 0.6 - 0.7% [19/20,22,24,25]. Cohorts studies of radiographic reports (3 studies; 2981 participants) from patients referred from primary care reported prevalences between 0.2% and 7.0% [28,31,34]. One retrospective cohort study in tertiary care (n = 1109 participants) found a prevalence of 5.9% [32].

A narrative review reported that most spinal malignancies were the result of metastases of primary tumours in the body, mainly from breast, lung or prostate cancer [18]. The review reported that just 20% of metastatic cancers appear in the lumbar spine; however, there was no reference to support this statement [18]. However, a descriptive study of 322 patients with cancer showed that most patients had lumbar metastasis, and in 32% of patients the primary cancer was breast cancer [27].

#### 3.3. Origin of red flags

The earliest report on red flags we located was by Thomas D Luke who reported red flags for spinal malignancies in 1911 [36]. He described 3 cases of women (age 50-61) with night pain and difficulties in walking/getting out of bed. Two women previously had a malignant breast cancer and all 3 women were in a “running down condition”, and died within a few weeks or months after first consultation. His final

recommendation was: “If there is severe pain at night running around the intercostal nerves to the front with inability to turn and a giving way of one or other leg on standing, one would be inclined to suspect spinal new growth. The previous existence of a breast tumour of malignant character is of great help . . . otherwise the difficulty of making an accurate diagnosis is undoubtedly very much greater.” [36]. This report is potentially the origin of the red flags ‘night pain’ and ‘previous history of cancer’.

We found two case series reporting red flags related to malignancy. One presented 14 cases of people referred to a tumour centre with pelvic cancer [37]. Based on the 14 patients the authors suggest to be alert when patients (age > 45 years) have ‘insidious onset of pain without a trauma’, ‘> one month duration of complaints’, ‘progressive pain’, ‘night pain’, ‘malaise’ and ‘anorexia’. Another study presented 2 patients who consulted a chiropractor for back pain in an outpatient facility [38]. Both patients received chiropractic care for 4 to 9 weeks and initially both patients responded positive to the treatment, but they returned within a couple of months with comparable complaints. Finally, both received extensive work-up (plain radiographs, MRI, CT-scans and blood work) when they were eventually diagnosed with malignancy. No red flags were apparent during the first visit.

We found five case reports. Patients were a 69-year-old man that visited a chiropractor with LBP after a twisting movement [39]; a 73-year-old lady with LBP and severe band like pain in the abdomen referred to a spine centre [40]; a 48-year-old women referred to physiotherapy for LBP [41]; a 15-year-old girl who visited the paediatric emergency department [42] and a 54-year-old man consulting the physiotherapist complaining of LBP after a motor vehicle accident [43]. In all cases, history taking and clinical examination did not reveal red flags that warranted medical referral. In addition, plain radiographs often showed no pathology, and treatment showed some benefit although for a short period. The main reason in all cases to order imaging and blood tests was progressive complaints within 2 to 6 weeks.

### 3.4. Studies without accuracy data

One study was a secondary analysis of randomised clinical trial data [23]. Patients with an episode of back pain were included and ‘familiarity of LBP’ was evaluated as an indicator for serious pathology. One (out of 1190) participant had a malignancy (metastasis of breast cancer), but it was unclear whether this participant was positive on the red flag or not.

Another study evaluated whether bone scintigraphy could identify occult malignancy in patients with musculoskeletal pain without a history of malignancy [28]. In this retrospective cohort of 491 patients, 52% of them had mid- or lower back pain. In total, 30 patients (out of 491) as having a malignancy; all

but 1 were older than 50 years. Of all 310 patients older than 50 years, 158 had complaints of mid- or lower back pain of which 13 had malignancy. Unfortunately, we were unable to calculate the false negative and true negative proportions in patients with LBP (younger than 50 years of age), so no accuracy data on 'age >50 years' are presented.

One study retrospectively evaluated patients referred to a spine centre to determine whether 'lack of pain' and 'no limitation in range of motion during the physical examination' were associated with malignancy [32]. They found 66 (out of 1109) patients to have metastatic bone cancer. Statistical testing revealed that these 66 patients weighted less, were more depressed or anxious, and more often had a scoliosis than the participants without malignancy ( $p < 0.05$ ). 'No pain during all four movements' showed a sensitivity of 55% (95%CI 42 - 67%) and a specificity of 59% (95%CI 56 - 62%). As the authors did not regard this a valid red flag nor is it endorsed in any guideline, it was not included in table 1. Another study aimed to identify risk factors for serious pathology in patients with non-traumatic LBP visiting an emergency department [33]. They retrospectively examined health records on serious pathology within 30 days after the visit. In a cohort of 329 patients with LBP, 22 cases of serious pathology were found. No data on malignancy separate were presented.

Recently, a study of health insurer data described the diagnosis of adolescents (10 – 19 years old) with LBP [35]. Within the PearlDriver Patient Record database, they found 215,592 adolescents with LBP and data of a final diagnosis within 1 year. In total 253 patients (0.1%) were diagnosed with a tumour over a 10-year period. No data on signs and symptoms are presented.

Lastly, a primary care prospective cohort included 669 patients with LBP of 55 years or older [25]. After a 12-months follow-up 4 patients were found with malignancy, but no data on the diagnostic accuracy of red flags for malignancy were reported.

### 3.5. Studies with accuracy data.

Deyo and Diehl undertook the earliest study found evaluating the diagnostic accuracy of possible red flags for malignancy in patients presenting with LBP in primary care [19/20]. At first, they established a prospective cohort with the aim to evaluate X-ray utilisation in patients with LBP ( $n=621$ ) visiting a walk-in clinic between March 1982 and May 1983 [19]. They assessed 65 history and physical examination items and collected data regarding use of X-ray imaging. Because they found some patients with malignancies, they extended this cohort and included patients with LBP until September 1984 ( $n= 1975$ ) [20]. Finally, 13 (0.7%) patients were identified as having malignancy 6 months after inclusion. Of these 13 cases, 10 were over 50 years old, 4 had a history of cancer, 3 suffered from unexplained weight loss,

and 6 cases had no specific signs or symptoms. Furthermore, there were 34 other patients who also had a history of cancer but no malignancy was found. The authors divided the patients into the ones that had a malignancy (n=13) and the ones without (n=1962). Four items from history taking appeared to be significantly more common (prevalent) in patients with malignancy compared to the ones that had not: 'age  $\geq$  50 years', 'previous history of cancer', 'sought medical care in the past month (not improving)' and 'duration of episode > 1 month'. Last, they developed a discriminant rule using multivariable analysis and 'a history of cancer' and an 'elevated ESR rate' correctly classifying 97% of the patients [20]. No accuracy data were presented on the discriminant rule.

In 1987, Frazier et al replicated this study in a retrospective cohort of 471 patients that visited a walk-in clinic with acute low back pain [29]. The authors stated: "The study was not intended to validate whether either the selective criteria or the clinic physicians' actual roentgenogram use correctly identified patients with serious causes of acute LBP." X-rays were taken for 99 patients at initial visit, which revealed one patient with a metastasis of the prostate.

Two more recent replication studies were conducted [24,25]. One prospective cohort study included 1172 patients with a new episode of LBP in primary care [24]. After a 12-months follow-up, none of the patients was diagnosed with a malignancy, although the 9 red flags evaluated had false positive rates ranging from 0.1% to 24.0%. For example, 202 patients (17.2%) had an 'insidious onset, 46 patients (3.9%) had 'a history of cancer', and 3 (0.3%) had 'unexplained weight loss'. The other study did not present accuracy data and is described above [25].

We found three other prospective cohort studies [21,22,26]. One study was performed at an emergency department, aiming to evaluate the reasons for clinicians to order spinal radiographs [21]. Of the 482 patients in the study, 458 (92%) reported LBP of which 7 patients (1.5%) were diagnosed with malignancy. All 7 patients were '60 years or older' and had 'a history of malignancy'. The clinicians were asked to state whether the patients had a 'suspected clinical diagnosis' of malignancy, but the diagnostic criteria for this clinical diagnosis are unclear. A malignancy was suspected in 20 patients but was confirmed by radiographs in only 3 of them.

The second study examined in a primary care cohort (n=1030) the diagnostic value of anteroposterior views on X-rays for serious pathology with the aim to 'safely omit it from examination without compromising diagnostic precision' [22]. In total, 886 (86%) patients had LBP as the main complaint. They found 1 case of malignancy on anteroposterior radiographs (it was unclear if the patient had LBP), whereas 20 patients were suspected to have a malignancy (suspected clinical diagnosis).

The third study was performed in secondary care setting aiming to describe the characteristic and MRI



findings of patients with LBP referred to an MRI center [26]. A total of 638 patients were included, of which 5 were diagnosed with a malignancy. Five red flags were evaluated: 'age > 50 years' (all 5 patients), 'continuous back pain' (0 patients), 'back pain at night' (2 patients), 'history of malignancy' (1 patient) and 'unexplained weight loss' (0 patients).

Finally, we found 2 systematic reviews evaluating the diagnostic accuracy of possible red flags for malignancy [13,15]. The first review was one of the first assessing the methodological quality of the included studies and could be considered a state-of-the-art review at that time [13]. This review included 8 studies including the landmark paper [20]. The other 7 studies evaluated red flags in patients with known cancer and their association with having spinal metastasis. Most of the studies did not explicitly mention (L)BP as main complaint. One study included 322 patients with known spinal metastasis in a hospital setting explicitly reported that in 36% of patients there was no accompanying LBP [27]. The review found 'sciatica' (history item) and 'spinal tenderness' (during physical examination) to be the most prominent factors related to the existence of vertebral metastasis, i.e. these are not red flags for malignancy in people with LBP. Spinal tenderness has been evaluated as a sign to detect vertebral cancer in hospital based patients with cancer varied in sensitivity between 54% and 80% (specificity 70%–78%), but this context is very different to screening for malignancy in primary care patients with LBP. When used to detect malignancy in patients with LBP in primary care its diagnostic accuracy is disappointingly low (sensitivity: 15%; specificity: 60%; LR+ <1) [13,20].

The second systematic review evaluated the diagnostic accuracy of all possible red flags as reported in 8 studies, including the cohort of Deyo & Diehl, but this cohort was regarded as 2 separate studies (1986 and 1988) [15]. All of the included studies are presented above.

### 3.6. Methodological quality

Three cohort studies in primary care were considered at low RoB (see figure 1) [20,24,26], 2 studies in secondary care were considered at moderate RoB, mainly because the reference standard was assessed with clear knowledge of the index tests results [21,22]. One retrospective cohort was regarded as of high RoB [29]. Only 1 study was clearly designed to assess diagnostic accuracy of red flags [24], all others were designed for another study question, but data about diagnostic accuracy could be extracted.

### 3.7. Diagnostic accuracy

Table 1 presents the 13 red flags as endorsed in the various guidelines and 2 additional red flags, of which the original authors regarded these flags as relevant.

For 5 red flags reported in one or more guidelines, we could not determine their origin or locate any diagnostic accuracy data (table 1). For the red flags not endorsed by the guidelines, one originated from the first prospective primary care cohort [20] and one from a case series [36].

Seven red flags were evaluated in more than 1 study, of which 2 were found clinically informative. They were 'history of malignancy' (moderate quality evidence; LR+ = 6.4-15.3) and 'strong clinical suspicion' (moderate quality evidence; LR+ = 12.0-54.2) The other 5 red flags we regarded as clinically uninformative.

#### **4. Discussion**

The prevalence of spinal malignancies in patients with LBP seemed to be influenced by setting and varies between 0% and 0.6% to 0.7% in primary care up to 0.2% and 7% in hospital settings. We found 13 red flags for malignancy endorsed in the 16 guidelines. For 5 of the red flags for malignancy the origin was unclear, and very limited diagnostic evidence was available for most guideline-endorsed red flags. The bulk of the evidence on red flags is provided by 2 studies considering a total of 3147 primary care patients [20,24]. Both studies were at low RoB and so can be considered as the best evidence on diagnostic accuracy of red flags in patients with LBP in primary care.

Two red flags ('history of malignancy' and 'strong clinical suspicion') show potential in identifying patients with a higher likelihood of malignancy as they have sufficiently high LR+. Five other red flags ('unexplained/unintentional weight loss', 'atypical pain', 'older age', 'failure to improve with treatment' and 'duration of complaint > 1 month') have quite modest LR+ and would not assist clinical decision making, given the low prevalence of malignancy in primary care. The fact that 1 red flag ('duration of complaint > 1 month') is not endorsed in any of the guidelines appears justified by the accuracy data. We were unable to draw conclusions on 7 red flags, of which 5 we were unable to find any information at all.

Atypical pain was a red flag with a range of definitions (including 'increasing pain at night', 'pain not eased by a prone position', 'continuous pain at rest', 'thoracic pain' and 'pain not worsened during movement') was endorsed in 12 guidelines. The first mention of the importance of pain was the study by Thomas Luke [36]. This study stated that "pain at night running around the intercostal nerves to the front", which might be the origin of 'thoracic pain' as a red flag in some of the guidelines. Actually, the factors stated by Thomas Luke in 1911 are endorsed by most guidelines. We found just 1 study from secondary care not supporting its use in clinical care.

#### 4.1. Comparison with other literature

Where the Cochrane review had strict selection criteria (diagnostic accuracy studies presenting data), our study selection was broader. We aimed to get a broader view by also exploring the origin of red flags and all kinds of supporting evidence to justify guideline recommendations. Our study supports the conclusions of the Cochrane review, hereby restating that the evidence for the use and recommendations of most red flags is very weak.

Most guidelines infer that when red flags are present further investigation is warranted [8]. This advice is nevertheless heavily criticized for the high risk of a false positive finding [9]. In a primary care cohort of 1172 patients with LBP, most patients (80.4%) had at least 1 red flag [24]. Also in a study focussing on physiotherapist's adherence to guidelines, the authors found that 7 of the 11 red flags were documented in 98% of the patients [44]. A web-based survey among adults with LBP confirmed these findings with 68.2% of the participants presenting at least 1 red flag [45].

As most patients have at least 1 individual red flag present, and most individual red flags do not show acceptable high accuracy, an alternative strategy could be to focus on a combination of red flags in identifying patients with a higher likelihood of malignancies (ruling people in). The first attempt to develop such a risk model is made by Deyo & Diehl [20]. They found that a combination of 'ESR rate > 20mm/hour' and 'a history of cancer' was both significant in a multivariate analysis, but this model still misclassified 6 out of 13 patients as having no malignancy while it was present.

One study developed a hypothetical model for diagnosing malignancy in primary care patients with LBP based on decision analysis with incremental cost-effectiveness analysis using the 5 most relevant factors found by Deyo & Diehl [46]. The strategy with the highest specificity and lowest costs was to refer all people who have one or more of the following clinical findings ('history of cancer', 'age > 50 years', 'weight loss', 'failure to improve with treatment') in combination with 'an ESR rate > 50mm/hour' or 'a positive X-ray' [46]. Unfortunately, this strategy still missed 2 patients with malignancy [Deyo 1988].

#### 4.2. Strength and limitations

The search strategy for this study was relatively broad, and we also included snowballing techniques, which meant that we started with reference and citation tracking of a land mark article, which is regarded a powerful strategy especially for finding papers in obscure locations [11]. Using this technique, we mainly found extra case studies. It is reassuring that all relevant studies, including diagnostic accuracy data, were found with the normal search strategy. Nevertheless, as most studies

were not designed as diagnostic accuracy studies, some potentially relevant studies may have been missed. In addition, data extraction was challenging and resulted in a few occasions in a different interpretation of the data compared to the systematic review [15]. Nevertheless, differences in data interpretation were small and did not affect our conclusions.

We considered the low prevalence of malignancy in primary care patients with LBP in setting the threshold of an LR+ >15 as potentially clinically informative red flags. This means that when the red flag is present the probability of malignancy increases from 0.8% to around 10%, which we considered a reasonable increase, given the nature of the target condition.

As diagnostic accuracy shows robust outcomes between various studies, we think our outcomes are generalizable to all different settings and populations.

The lack of a precise definition of red flags highlights a barrier when determining which red flags may be clinically informative. For example, because of lack of definition, conflicting diagnostic accuracy data may be presented for the same red flag in similar clinical settings (e.g. strong clinical suspicion'). Also, concerning the red flag 'history of malignancy', the timing and clinical characteristics of previous malignancy might be important. The origin of this red flag comes from Thomas Luke in 1911 based on women with a history of breast cancer. None of the other studies provided clinical information on previous malignancy or the time elapsed between diagnosis of the previous malignancy and the current episode of back pain, which might influence the accuracy of this red flag. Similarly, potential overlap of red flags might be an additional barrier, as it is unclear whether 'increasing pain at night', pain not eased by a prone position' and 'continuous pain at rest' are similar.

#### 4.3. Future direction

Malignancy is a disease clinicians do not want to miss, but equally a clinician does not want to needlessly alarm their patients by sending patients off for a diagnostic work-up. Previous discussions on screening for malignancy in patients presenting with LBP have tended to promote the ideal of a highly sensitive test where a negative result can be used to 'rule out' malignancy. The limitations with this approach are that the pretest probability of malignancy is already very low and the use of a highly sensitive test will inevitably mean that most people who test positive will be false positives. The American College of Physicians Guideline on imaging for LBP offers an alternative approach. They suggest that when there is a weaker suspicion of malignancy, the clinician could consider initiating a trial of therapy and reviewing the patient in a timely fashion to ensure that recovery is occurring as expected. At present, we are

unaware of any research that has surveyed clinicians and patients about their views and preferences in this matter. This would seem an important area for future research.

In addition, we suggest a (limited) core set of red flags, which could be evaluated in new high quality diagnostic accuracy studies and to start with a core set of 2 red flags that show to be promising in the current study. Developing a diagnostic (risk) model might result in better diagnostic accuracy.

## **5. Conclusion**

For the majority of red flags for malignancy included in clinical guidelines, the origin is unclear and there is strikingly little or no evidence available regarding their diagnostic accuracy. Two red flags were evaluated in 2 diagnostic studies at low RoB and had acceptably high LR+ to guide decision making: 'history of malignancy' and 'strong clinical suspicion'. At present, these are the only red flags that have empirical basis for inclusion in clinical guidelines.

## Reference

1. Jones LD, Pandit H, Lavy C. Back pain in the elderly: a review. *Maturitas*. 2014;78(4):258-62.
2. Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, York J, Das A, McAuley JH. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. *BMJ* 2008;337:a171.
3. Tulder M v, Becker A, Bekkering T, Breen A, Gil del Real MT, Hutchinson A. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 2006;15 (Suppl. 2): S169–S191.
4. Raison NT, Alwan W, Abbot A, Farook M, Khaleel A. The reliability of red flags in spinal cord compression. *Arch Trauma Res*. 2014 Mar 30;3(1):e17850.
5. Jarvik JG, Deyo RA. Diagnostic evaluation of low back pain with emphasis on imaging. *Ann Intern Med*. 2002;137(7):586–97.
6. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA* 1992;268(6):760-5.
7. Koes B, Tulder M, Wei C, Lin C, Macedo G, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J* 2010;19:2075–2094.
8. Chou R, Qaseem A, Snow V, Casey D, Cross JT Jr, Shekelle P, Owens DK; Clinical Efficacy Assessment Subcommittee of the American College of Physicians; American College of Physicians; American Pain Society Low Back Pain Guidelines Panel. 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):478-91.
9. Underwood M, Buchbinder R. Red flags for back pain. *BMJ*. 2013;347:f7432.
10. Verhagen AP, Downie A, Popal N, Maher C, Koes BW. Red flags presented in current low back pain guidelines: a review. *Eur Spine J* 2016. Epub ahead July 2016.
11. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 2005;331(7524):1064-1065.
12. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-36.
13. Van den Hoogen HJ, Koes BW, Van Eijk JT, Bouter LM. On the diagnostic 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;20:318-327.
14. Henschke N, Maher CG, Refshauge KM. Screening for malignancy in low back pain patients: a systematic review. *Eur Spine J*. 2007;16(10):1673-9.
15. Henschke N, Maher CG, Ostelo RW, de Vet HC, Macaskill P, Irwig L. Red flags to screen for malignancy in patients with low-back pain. *Cochrane Database Syst Rev* 2013;2:CD008686

16. Downie A, Williams CM, Henschke N, Hancock MJ, Ostelo RW, de Vet HC, Macaskill P, Irwig L, van Tulder MW, Koes BW, Maher CG. Red flags to screen for malignancy and fracture in patients with low back pain: systematic review. *BMJ*. 2013;347:f7095.
17. Downie A, Williams CM, Henschke N, Hancock MJ, Ostelo RW, de Vet HC, Macaskill P, Irwig L, van Tulder MW, Koes BW, Maher CG. Red flags to screen for malignancy and fracture in patients with low back pain. *Br J Sports Med*. 2014;48(20):1518.
18. Van Goethem JW, van den Hauwe L, Ozsarlak O, De Schepper AM, Parizel PM. Spinal tumors. *Eur J Radiol*. 2004;50(2):159-76.
19. Deyo RA, Diehl AK. Lumbar spine films in primary care: current use and effects of selective ordering criteria. *J Gen Intern Med* 1986;1:20-5
20. Deyo RA, Diehl AK. Cancer as a cause of back pain: frequency, clinical presentation, and diagnostic strategies. *J Gen Intern Med* 1988;3:230-8.
21. Reinus WR, Strome G, Zwemer FL Jr. Use of lumbosacral spine radiographs in a level II emergency department. *AJR Am J Roentgenol* 1998;170:443-7.
22. Khoo LA, Heron C, Patel U, Given-Wilson R, Grundy A, Khaw KT, Dundas D. The diagnostic contribution of the frontal lumbar spine radiograph in community referred low back pain--a prospective study of 1030 patients. *Clin Radiol*. 2003;58(8):606-9.
23. Donner-Banzhoff N, Roth T, Sönnichsen AC, Luckmann J, Leonhardt C, Chenot JF, et al. Evaluating the accuracy of a simple heuristic to identify serious causes of low back pain. *Family Practice* 2006;23(6):682–6.
24. Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, York J, Das A, McAuley JH. 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):3072-80.
25. Enthoven WT, Geuze J, Scheele J, Bierma-Zeinstra SM, Bueving HJ, Bohnen AM, Peul WC, van Tulder MW, Berger MY, Koes BW, Luijsterburg PA. Prevalence and "Red Flags" Regarding Specified Causes of Back Pain in Older Adults Presenting in General Practice. *Phys Ther*. 2016;96(3):305-12.
26. de Schepper EI, Koes BW, Veldhuizen EF, Oei EH, Bierma-Zeinstra SM, Luijsterburg PA. Prevalence of spinal pathology in patients presenting for lumbar MRI as referred from general practice. *Fam Pract*. 2016;33(1):51-6.
27. Schaberg J, Gainor BJ. A profile of metastatic carcinoma of the spine. *Spine* 1985;10:19-20
28. Jacobson AF. Musculoskeletal pain as an indicator of occult malignancy: yield of bone scintigraphy. *Archives of Internal Medicine* 1997;157:105–9.
29. Frazier LM, Carey TS, LylesMF, Khayrallah MA, McGaghie WC. Selective criteria may increase lumbosacral spine roentgenogram use in acute low-back pain. *Archives of Internal Medicine* 1989;149:47–50.
30. Hollingworth W, Todd CJ, King H, Males T, Dixon AK, Karia KR, Kinmonth AL. Primary care referrals for lumbar spine radiography: diagnostic yield and clinical guidelines. *Br J Gen Pract*. 2002;52(479):475-80.

31. Van den Bosch MAAJ, Hollingworth W, Kinmonth AL, Dixon AK. Evidence against the use of lumbar spine radiography for low back pain. *Clin Radiol* 2004;59:69-76.
32. Cook C, Ross MD, Isaacs R, Hegedus E. Investigation of nonmechanical findings during spinal movement screening for identifying and/or ruling out metastatic cancer. *Pain Pract* 2012;12:426-33.
33. Thiruganasambandamoorthy V, Turko E, Ansell D, Vaidyanathan A, Wells GA, Stiell IG. Risk factors for serious underlying pathology in adult emergency department nontraumatic low back pain patients. *J Emerg Med*. 2014;47(1):1-11.
34. Leichtle UG, Wünschel M, Socci M, Kurze C, Niemeyer T, Leichtle CI. Spine radiography in the evaluation of back and neck pain in an orthopaedic emergency clinic. *J Back Musculoskelet Rehabil*. 2015;28(1):43-8.
35. Yang S, Werner BC, Singla A, Abel MF. Low Back Pain in Adolescents: A 1-Year Analysis of Eventual Diagnoses. *J Pediatr Orthop*. 2015 Sep 11. [Epub ahead of print]
36. Luke TD. Three cases of malignant disease of the spine. *BMJ* 1911;april 29:988-99-89
37. Thompson RC Jr, Berg TL. Primary bone tumors of the pelvis presenting as spinal disease. *Orthopedics*. 1996;19(12):1011-6.
38. Pringle RK, Wyatt LH. The appropriate use of radiography in clinical practice: a report of two cases of biomechanical versus malignant spine pain. *Chiropr Osteopat*. 2006;14:8.
39. Rectenwald R. A case study of back pain and renal cell carcinoma. *J Chiropr Med*. 2008;7(1):24-7.
40. Finucane L. Metastatic disease masquerading as mechanical low back pain; atypical symptoms which may raise suspicion. *Man Ther*. 2013;18(6):624-7.
41. Mabry LM, Ross MD, Tonarelli JM. Metastatic cancer mimicking mechanical low back pain: a case report. *J Man Manip Ther*. 2014;22(3):162-9.
42. Massoud M, Del Bufalo F, Caterina Musolino AM, Schingo PM, Gaspari S, Pisani M, Orazi C, Reale A, Raucci U. Myeloid Sarcoma Presenting as Low Back Pain in the Pediatric Emergency Department. *J Emerg Med*. 2016 Jul 6. [Epub ahead of print]
43. Plass LM, McGee TG, Elliott JM. Identification of Metastatic Lesions in a Patient With Low Back Pain Following a Motor Vehicle Collision. *J Orthop Sports Phys Ther*. 2016;46(2):124.
44. Leerar PJ, Boissonnault W, Domholdt E, Roddey T. Documentation of red flags by physical therapists for patients with low back pain. *J Man Manip Ther*. 2007;15(1):42-9.
45. Miekisiak G, Latka D, Sobolewski T, Sulewski A, Kubaszewski L, Kaczmarczyk J. Prevalence of "red flags" among the internet user with low back pain seeking information on spinal health-a web survey. *Telemed J E Health*. 2015;21(2):120-4.
46. Joines JD, McNutt RA, Carey TS, Deyo RA, Rouhani R. Finding cancer in primary care outpatients with low back pain: a comparison of diagnostic strategies. *J Gen Intern Med*. 2001;16(1):14-23.



**Table 1: Red flags for malignancies**

<b>Red flags Endorsed in number of guidelines)</b>	<b>Prevalence</b>	<b>Setting</b>	<b>Identified as relevant in cohort studies or case series</b>	<b>Diagnostic accuracy</b>		<b>Risk of Bias</b>	<b>Quality of the evidence (informativeness)</b>
				<b>Se (95%CI)</b>	<b>Sp (95%CI)</b>		
History of malignancies/cancer (15)	N/A	Unknown	Case series [Luke 1911, Thompson 1996]	N/A			Moderate (informative)
	0.66% (13/1975)	Primary care	Prospective cohort [Deyo 1988]	31% (6 to 56)	98% (97 to 99)	Low risk	
	0% (0/1172)	Primary care	Prospective cohort [Henschke 2009]		96% (95 to 97)	Low risk	
	0.7% (5/638)	Secondary care	Prospective cohort [de Schepper 2016]	20% (-15 to 55)	97% (96 to 98)	Low risk	
	1.45% (7/482)	Emergency department	Prospective cohort [Reinus 1998]*	100%		Moderate risk	
(Unexplained/unintentional) Weight loss (12)	0.66% (13/1975)	Primary care	Prospective cohort [Deyo 1988]	15% (-4 to 35)	94% (93 to 95)	Low risk	Moderate (uninformative)
	0% (0/1172)	Primary care	Prospective cohort [Henschke 2009]		99.7% (99 to 100)	Low risk	
	0.7% (5/638)	Secondary care	Prospective cohort [de Schepper 2016]	0% (0 to 0)	88% (86 to 91)	Low risk	
<b>Pain</b> either increasing at night or at rest or pain at night that is not eased by a prone position (or increasing in supine position) (12)	N/A	Unknown	Case series [Luke 1911, Thompson 1996]	N/A			Very low (uninformative)
	0.7% (5/638)	Secondary care	Prospective cohort [de Schepper 2016] (low back pain at night)	40% (-3 to 83)	47% (43 to 50)	Low risk	
<b>Older age;</b> either just older age or more specific over 50 years (7)	N/A	Tertiary care	Case series [Thompson 1996]				Very low (uninformative)
	0.66% (13/1975)	Primary care	Prospective cohort [Deyo 1988] (age > 50 years)	77% (54 to 100)	71% (69 to 73)	Low risk	
	0% (0/1172)	Primary care	Prospective cohort [Henschke 2013] (age > 55)		66% (63 to 69)	Low risk	
	0.7% (5/638)	Secondary care	Prospective cohort [de Schepper 2016] (age > 50 years)	100%	77% (74 to 80)	Low risk	
	1.45% (7/482)	Emergency department	Prospective cohort [Reinus 1998]*, (age > 50	100%		Moderate risk	

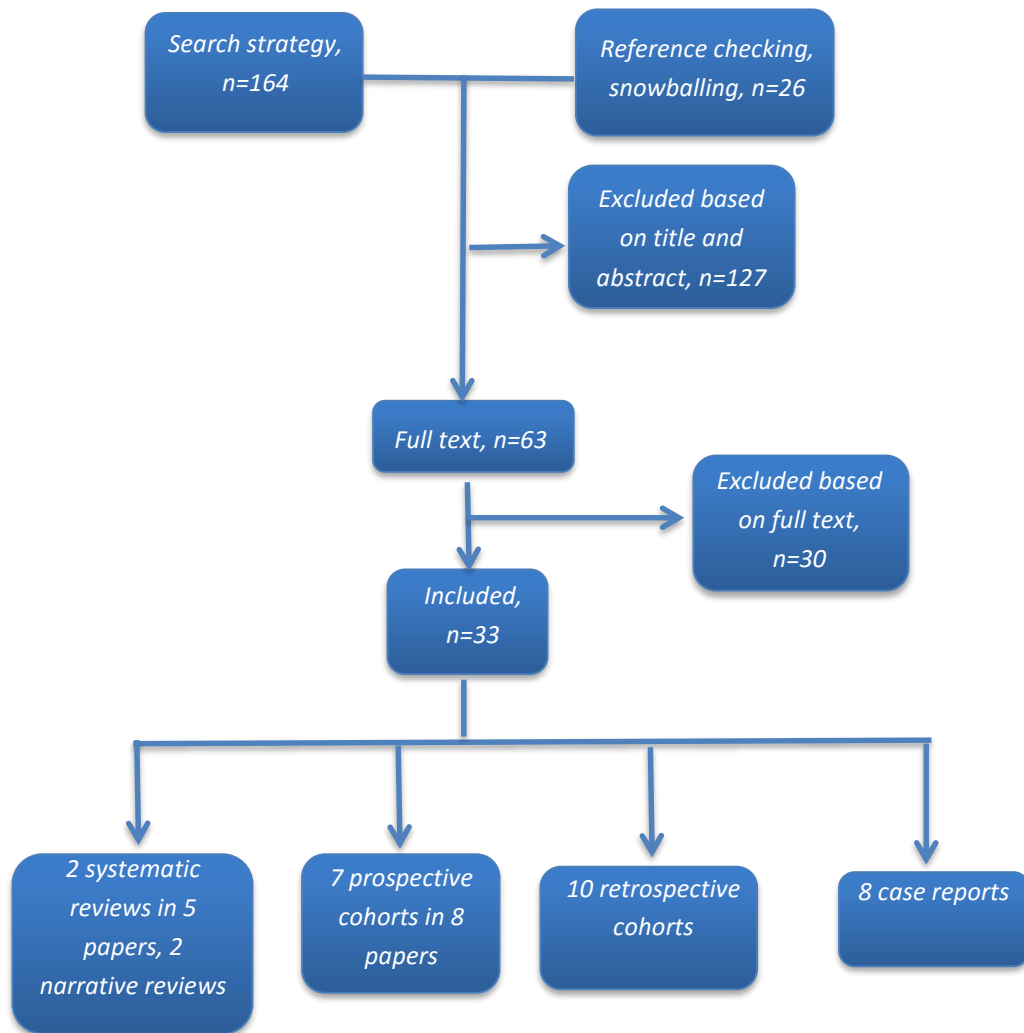
<b>Red flags Endorsed in number of guidelines)</b>	<b>Prevalence</b>	<b>Setting</b>	<b>Identified as relevant in cohort studies or case series</b>	<b>Diagnostic accuracy</b>		<b>Risk of Bias</b>	<b>Quality of the evidence (informativeness)</b>
				<b>Se (95%CI)</b>	<b>Sp (95%CI)</b>		
			years)				
	0.45% (2/471)	Secondary care	Retrospective cohort [Frazier 1989], (age > 50 years)	50% (-19 to 119%)	74% (70 to 78)	High risk	
Malaise (5)	N/A	Tertiary care	Case series [Thompson 1996]				NA
Failure to improve with treatment (>4 – 6 weeks) / Seeking medical care last month (4)	N/A		Case study [Pringle 2006, Mabry 2014]	N/A			Moderate (uninformative)
	0.66% (13/1975)	Primary care	Prospective cohort [Deyo 1988]	31% (6 to 56)	90% (89 to 91)	Low risk	
	0% (0/1172)	Primary care	Prospective cohort [Henschke 2009]		84% (81 to 91)	Low risk	
Strong clinical suspicion (1)	(1/1030)	Primary care	Prospective cohort [Khoo 2003]	100%	98% (97 to 99)	Moderate risk	Moderate (informative)
	0% (0/1172)	Primary care	Prospective cohort [Henschke 2009]		99.9% (99.9 to 100)	Low risk	
	1.45% (7/482)	Emergency department	Prospective cohort [Reinus 1998]	43% (6 to 80)	96% (95 to 98)	Moderate risk	
Fever (1)	0.66% (13/1975)		Prospective cohort [Deyo 1988]		98% (97 to 99)	Low risk	NA
Reduced appetite (1)							
Rapid fatigue (1)							
Progressive symptoms (1)							
<b>Red flags not endorsed in guidelines</b>							
Multiple risk factors (1)							
Paraparesis (1)				)			
Duration of the complaint > 1 month (0)	N/A	Tertiary care	Case series [Thompson 1996]				Moderate (uninformative)
	0.66% (13/1975)	Primary care	Prospective cohort [Deyo 1988]	50% (22 to 78)	81% (79 to 83)	Low risk	

<b>Red flags Endorsed in number of guidelines)</b>	<b>Prevalence</b>	<b>Setting</b>	<b>Identified as relevant in cohort studies or case series</b>	<b>Diagnostic accuracy</b>		<b>Risk of Bias</b>	<b>Quality of the evidence (informativeness)</b>
				<b>Se (95%CI)</b>	<b>Sp (95%CI)</b>		
	0% (0/1172)	Primary care	Prospective cohort [Henschke 2009]		93% (91 to 94)	Low risk	




Se = sensitivity; sp = specificity; 95%CI = 95% confidence interval

N/A = not applicable as study only included cases so prevalence could not be calculated

**Figure 1: Study flow chart**



**Figure2: Risk of Bias assessment**

	<u>Risk of Bias</u>				<u>Applicability Concerns</u>		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
de Schepper 2016	+	+	+	?	+	+	+
Deyo 1988	+	?	?	+	+	+	+
Frazier 1987	?	-	-	-	?	?	?
Henschke 2009	+	+	?	+	+	+	+
Khoo 2003	+	+	-	+	+	+	+
Reinus 1998	+	+	-	+	?	+	?
<div>  <b>High</b>  <b>Unclear</b>  <b>Low</b> </div>							