INTRODUCTION

When a patient consults a GP, part of the management of the individual involves an assessment of the probability of serious illness including malignancy. One way of expressing this probability is in terms of the positive predictive value (PPV): the proportion of people with the symptom, sign, or test result who develop cancer. Research on medical decision models reveals that probabilities and the valuation of outcomes are distorted by individuals,\(^1\) and thus theories of behaviour based on cognitive assessment of risk are not directly applicable to the consultation.\(^2\) For instance, in one model of the consultation, the ‘evidence’ from research that allows the calculation of probabilities of risk and benefit is modified by an individual doctor’s knowledge, skills, attitudes, resources, and legal requirements, and an individual patient’s ideas, concerns, expectations, beliefs, and values, to produce the final outcome.\(^3\) These factors vary widely between individuals.

Decision makers who wish to maximise desirable outcome often adopt models based on probabilities of benefits and harms. Predictive values have been incorporated into guidelines for primary care published by various colleges, academies, and institutes, including the National Institute for Health and Clinical Excellence,\(^4\) concerning who to investigate or refer for possible malignancy in primary care. A major criticism
is that much of the evidence that underpins these guidelines is not derived from those patients who consult a GP during a routine consultation. This is important, because predictive values are dependent on the prevalence of the disease in the population from which they are derived and thus applied. The prevalence of cancers within populations and their associated PPVs of symptoms, signs, and non-diagnostic test results increase from community to primary care to hospital populations.

The level of the predictive value given in guidelines at which action is required to investigate or not for cancer will depend on the consequences of a delay in the diagnosis of the disease and the physical, psychological, and material cost of investigation. No guidelines for cancer have been published that state the level of risk that necessitates action. NICE guidelines for the primary prevention of cardiovascular disease recommend statin therapy for hyperlipidaemia when the 10-year cardiovascular risk is 20% or more. This threshold for initiating statin therapy was based on a cost-effectiveness analysis and risk of adverse effects. Studies show a reduction in relative risk of 27% (95% confidence interval [CI] = 14 to 37%) from statin therapy in patients without cardiovascular disease. This equates to an absolute risk reduction of approximately 5% or a number needed to treat with statins to prevent a cardiovascular event of 20. Statins are a safe group of drugs with a low risk of serious adverse effects at standard doses (less than 1 in 10 000 patient-years for myopathy). The consequences of delayed cancer diagnosis in terms of morbidity and mortality are comparable to those of cardiovascular disease, and generally the risks of cancer investigation are similar to the risk of statin therapy. For instance, the rate of perforation of the bowel for colonoscopy used in the investigation for colorectal cancer is 1:769 and procedure-related mortality 1:1537.

Fear of cancer may mean it is perceived by patients as a greater risk than cardiovascular disease, even though a greater number of people die of the latter. Emotional reactions to risky situations often diverge from cognitive assessments of those risks. While the level of risk an individual should take cannot be defined, comparison of risk of cancer to that of cardiovascular disease suggests that if a population consults with a symptom that could be cancer, then the maximum acceptable risk of not excluding cancer is 1 in 20, which equates to a PPV of 5%.

The objective of this systematic review was to identify the symptoms, signs, and non-diagnostic test results in people routinely consulting a GP that have a PPV of 5% or more for cancer and hence mandate action by a GP except in individual patient-centred circumstances.

**How this fits in**

The positive predictive values of symptoms, signs, and non-diagnostic test results for cancer in primary care are generally low. The level of risk for cancer that a patient and doctor are willing to accept varies. Using systematic review, good evidence has been gathered for eight symptoms, signs, and non-diagnostics test results as being highly predictive of cancer in primary care. Recommendations for clinical practice, audit, education, and research are made.

**METHOD**

**Search strategy**

Fourteen electronic databases were searched from their commencement to October 2009 (Box 1), using terms for primary care, cancer, and predictive values. Searches were limited to English language and human studies. The search strategy used in MEDLINE is shown in Box 2.

Reference lists of relevant studies were checked and citations were tracked using electronic databases. Experts in the field were emailed for further information. All the references were downloaded into reference management software, and the duplicates were removed.

**Inclusion and exclusion criteria**

Primary studies on unselected primary care populations (patients who consult a GP during a routine consultation) that examined symptoms, signs, or tests for a possible cancer were included in the review. The study had to report a PPV for the symptom, sign, or test, or contain the information necessary to calculate it. For final inclusion, the presented or

**Box 1. Electronic databases searched.**

- MEDLINE (from 1950)
- EMBASE (1980)
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (1981)
- Web of Science (1970)
- Biosis Previews (1969)
- Ageline (1978)
- Applied Social Sciences Index and Abstracts (1987)
- Social Services Abstracts (1979)
- Sociological Abstracts (1952)
- Academic Search Elite (1975)
- British Nursing Index (2004)
- International Bibliography of the Social Sciences (1919)
- Health Management Information Consortium (1919)
Box 2. Search (example taken from search used for the MEDLINE database).

"**GP**" OR exp† FAMILY PRACTICE OR exp. PHYSICIANS, FAMILY OR exp. PRIMARY HEALTH CARE OR (family ADJ2 pract*) OR (general ADJ2 pract*) OR (primary ADJ2 care) OR (family ADJ (physician OR doctor*))

AND

"exp. NEOPLASMS OR neoplasm* OR cancer* OR carcinoma* OR malignan* OR lesion* OR tumo?§"

AND

"(risk* ADJ3 cancer) OR "PPV" OR (predictive ADJ value*) OR (diagnos* ADJ value*) OR (early ADJ diagnosis) OR (conditional ADJ probabilit*) OR sensitivity OR specificity OR predictor* OR prevalen* OR (time ADJ interval) OR inciden* OR exp "REFERRAL AND CONSULTATION" OR alarm features"

* truncation — picks up various different word endings. † exp. = exploded thesaurus term; picks up all articles which contain the particular terms as a keyword. ‡ ADJ = adjacent; picks up words which have up to 3 words in between them. § ? = picks up alternative spellings.

calculated PPV had to be:

- 5% or more either for the whole study population or for an age/sex stratum; or
- less than 5% but for the same symptom, sign, or test another eligible study gave a point estimate for the PPV of 5% or more.

The latter criterion was necessary in order to calculate a pooled PPV using all the evidence for a symptom, sign, or test and to compare studies with conflicting results. Exclusion criteria for the review were:

- screening studies, as these involve asymptomatic individuals;
- tests in which a histological diagnosis could be made, as clinical judgement on the need for further investigation or referral would be removed (for example, endoscopy with biopsy, skin biopsy);
- metastatic, recurrent, or cancer-therapy studies as these are not early disease;
- studies with 15 cases or fewer, as these are unlikely to produce reliable conclusions;
- studies written in non-English language due to lack of resources for translating papers; and
- studies in healthcare systems that are not comparable to the UK.

**Study selection and assessment of methodological quality**

One reviewer screened titles and abstracts and retrieved papers for inclusion. A second reviewer checked 150 (7%) of the abstracts and 25% of the full papers to ensure they met the eligibility criteria. All studies identified in the screening of titles and abstracts that met the eligibility criteria, or for which it was not possible to tell, were retrieved in full. Two authors independently carried out an unblinded assessment of the risk of bias in each eligible study, using the Newcastle-Ottawa Quality Assessment Scales (NOQAS), for cohort and case-control studies. Studies are rated from 1 to 9 stars in the NOQAS, with 9 stars indicating a high-quality study. Any discrepancies or disagreements were discussed, and if consensus could not be achieved this was resolved by a third reviewer.

**Data extraction and analysis**

Once the final selection of articles for inclusion had been agreed, two researchers independently extracted data using a standardised data-extraction form. This included the type of study, characteristics of the study participants, the duration of follow-up, and the reported or calculated PPV.

Any subgroups (age and/or sex) reported in the studies that contained fewer than 15 people, or where authors stated that numbers were too low to reasonably calculate a CI and could not be combined with other data, were excluded from the review.

In studies that did not report the PPV, it was calculated as the proportion of those with that symptom, sign, or test who developed cancer. Where 95% CIs for the PPV were not stated in the paper, these were estimated, where possible, using the data presented. Where this was not possible, an attempt was made to contact the authors.

A meta-analysis was performed to obtain a pooled PPV for symptoms reported in four or more studies. The \( I^2 \) statistic was first calculated to assess the heterogeneity of the studies examining a symptom. The \( I^2 \) statistic can be interpreted as the proportion of total variation in study estimates that is due to heterogeneity between studies. For risk factors where studies could be considered homogenous, a meta-analysis using a fixed-effects (no significant interstudy variation) approach was used. Otherwise, a random-effects approach (significant interstudy variation) based on the inverse variance method was used.

The analysis was performed using Stata (version 10) and confidence interval analysis software.

**RESULTS**

The search resulted in 17 889 unique references. Of these, 25 papers were found to meet all the eligibility criteria (Figure 1).

Sixteen of the included studies were conducted in the UK, two in each of the Netherlands, Belgium, and US, and one in each of Australia, Denmark, and Germany.

Recruitment into the studies was by one of three
processes: retrospective analysis of clinical records, recruitment of patients with symptoms by GPs either as part of routinely collected data or specifically for the study, and postal survey with prospective identification of cancer from GP records.

Studies varied in the method of extraction of medical record data. The source of medical record data was either routinely coded data only, all routinely recorded data including free text, or GP registered onto specific doctor-completed forms. The duration of follow-up when it was reported ranged from no follow-up to 11 years. Five studies reported that cancer was excluded prior to the first consultation.

Nineteen of the studies had a score of ≥6 on the NOQAS, indicating good quality. Common flaws were: the cohort was unrepresentative of unselected primary care populations; differences in criteria determining the presence of cancer; incomplete or inadequate blinding of assessment; lack of demonstration that the individual had not had the relevant cancer diagnosed previously; short duration of follow-up; and loss to follow-up.

**Cancer site**

**Gastrointestinal.** For colorectal cancer, 12 studies were identified giving a predictive value of 5% or more, and three further studies were identified giving a value of less than 5% for one of the symptoms or combinations of symptoms. In 11 of the studies, rectal bleeding (definitions varied) had a predictive value of 5% or more in at least one of the age–sex stratifications (Table 1).

Eight of these had quality scores of ≥6.

In seven of the 11 studies there was, however, potential selection bias as GPs recruited participants into the study. In three it was stated that participants may not have been representative of all those consulting with the symptom, and in two studies it was demonstrated that the effect of this potential bias was not significant. In three additional studies, rectal bleeding had a PPV of less than 5% in all the
subgroups presented (Table 2).20–22 One study identified patients from large electronic databases with a computer code of change in bowel habit in the medical record and used no free text. In a retrospective cohort study, Lawrenson et al reported a PPV of greater than 5% in men aged ≥60 years for colorectal cancer, while Hamilton and co-workers, in a case-control study, reported PPVs of less than 5% for all age groups (Table 3).24

In two prospective cohort studies, and a randomised controlled trial, the symptom of rectal bleeding was combined with other symptoms or test results to give PPVs of 5% or more (Table 4). However, the numbers of symptomatic patients were generally low and CIs wide. Any conclusions drawn from these findings are weak. An additional case-control study gave some combinations of symptoms and signs a

Table 1. Studies giving a PPV of ≥5% in one or more strata of rectal bleeding for colorectal cancer.

<table>
<thead>
<tr>
<th>Study and NOQAS star rating</th>
<th>Type of cohort</th>
<th>Data source</th>
<th>Follow-up</th>
<th>Age, years</th>
<th>Sex</th>
<th>Symptom</th>
<th>Study</th>
<th>NOQAS Data Number with symptom</th>
<th>PPV, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones et al, Retrospective</td>
<td>Routinely recorded computer codes</td>
<td>3 years</td>
<td>≥15</td>
<td>All</td>
<td>Rectal bleeding computer codes</td>
<td>M</td>
<td>75–84</td>
<td>633</td>
<td>7.7 (5.8 to 10.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td>≥85</td>
<td>157</td>
<td>5.1 (2.2 to 9.8)</td>
</tr>
<tr>
<td>Parker et al, Retrospective</td>
<td>Routinely recorded computer codes</td>
<td>2 years</td>
<td>≥25</td>
<td>All</td>
<td>First ever consultation with rectal bleeding computer code</td>
<td>M+F</td>
<td>75–84</td>
<td>Number with cancer 173</td>
<td>5.5 (4.7 to 6.3)</td>
</tr>
<tr>
<td>Lawrenson et al, Retrospective</td>
<td>Routinely recorded computer codes</td>
<td>1 year</td>
<td>40–89</td>
<td>All</td>
<td>Rectal bleeding computer codes</td>
<td>M</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Waubers et al, Retrospective</td>
<td>Routinely registered</td>
<td>18–30 months</td>
<td>All</td>
<td>All</td>
<td>Cases of rectal bleeding</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Du Toit et al, Prospective</td>
<td>Specifically registered</td>
<td>Not given</td>
<td>≥45</td>
<td>All</td>
<td>New onset rectal bleeding</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Heintze et al, Prospective</td>
<td>Specifically registered</td>
<td>Not given</td>
<td>≥15</td>
<td>All</td>
<td>Cases rectal bleeding</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Ellis and Thompson, Prospective</td>
<td>Specifically registered</td>
<td>18 months</td>
<td>&gt;34</td>
<td>All</td>
<td>Primary complaint rectal bleeding</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Nørrelund and Nørrelund, Prospective</td>
<td>Specifically registered</td>
<td>22–57 months</td>
<td>≥40</td>
<td>All</td>
<td>First episode or change in rectal bleeding pattern</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Metcalf et al, Prospective</td>
<td>Specifically registered</td>
<td>Not given</td>
<td>≥40</td>
<td>All</td>
<td>Rectal bleeding of recent onset</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Mant et al, Prospective</td>
<td>Specifically registered</td>
<td>Not given</td>
<td>≥40</td>
<td>All</td>
<td>Rectal bleeding &lt;6 months onset</td>
<td>M+F</td>
<td>60–69</td>
<td>Not given</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70–79</td>
<td>Not given</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80–89</td>
<td>Not given</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Fijten et al, Prospective</td>
<td>Specifically registered</td>
<td>Mean 20 months SD 5 months</td>
<td>18–75</td>
<td>All</td>
<td>Overt rectal bleeding reason for encounter or history of rectal blood loss visible to the patient in the previous 3 months</td>
<td>M+F</td>
<td>60–75</td>
<td>Not given</td>
<td>6.0</td>
</tr>
</tbody>
</table>

*Not possible to calculate confidence intervals (author contacted, original data not available). F = female. M = male. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value. SD = standard deviation.
Table 2. Studies giving a PPV of <5% of rectal bleeding for colorectal cancer.

<table>
<thead>
<tr>
<th>Study and NOQAS star rating</th>
<th>Type</th>
<th>Data</th>
<th>Follow-up</th>
<th>Age, years</th>
<th>Sex</th>
<th>Number with symptom</th>
<th>Symptom</th>
<th>Symptom C omments</th>
<th>PPV (95% CI)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton et al, 2009* (8 stars)</td>
<td>Case-control</td>
<td>Routinely recorded computer codes</td>
<td>NA</td>
<td>≥30</td>
<td>M+F</td>
<td>853 cases, 460 controls</td>
<td>Rectal bleeding computer codes</td>
<td>PPV stratified by age and sex but numbers of cases and controls by age and sex not given</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamilton et al., 2005* (9 stars)</td>
<td>Case-control</td>
<td>Routinely recorded data including free text</td>
<td>NA</td>
<td>≥40</td>
<td>M+F</td>
<td>148 cases, 73 controls</td>
<td>Rectal bleeding</td>
<td>PPV 2.4% (95% CI = 1.9 to 3.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson et al., 1999* (3 stars)</td>
<td>Prospective cohort</td>
<td>Self-reported questionnaire</td>
<td>4–5 years</td>
<td>&gt;16</td>
<td>M+F</td>
<td>197</td>
<td>Rectal bleeding and consulted GP</td>
<td>No cases of rectal cancer; two cases of caecal cancer felt not to cause bleeding but not stated if in consultation cohort</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norrelund and Norrelund, 1996* (3 stars)</td>
<td>Prospective cohort</td>
<td>Specifically registered</td>
<td>22–57 months</td>
<td>≥40</td>
<td>M+F</td>
<td>45</td>
<td>Current bleeding episode similar to previous</td>
<td>PPV 4.4% (95% CI = 1.2 to 14.8)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F = female. M = male. NA = not applicable. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value.

PPV above 5%, but the authors did not calculate CIs for PPVs due to small sample sizes.23

One study was identified evaluating iron deficiency anaemia as a predictor for upper and lower gastrointestinal cancer;26 one for cancer;25 and one for colorectal cancer only (Table 5).25 The study by Yates et al identified 71 patients with cancer from a cohort of 431 people with iron deficiency anaemia.26 Of these, 35 had lower gastrointestinal cancer, 23 non-gastrointestinal cancer, and 13 other gastrointestinal cancer. The PPV for colorectal cancer was above 5% in men aged >20 years with a haemoglobin of ≤12 g/dl, and the PPV was less than 5% in women aged >50 years with a haemoglobin of ≤11 g/dl. Hamilton et al used a case-control methodology but felt unable to calculate a PPV for colorectal cancer in men aged 30 to 59 years, as sample sizes were small.26 Only in cohorts aged ≥60 years did the PPV become greater than 5%, although in some strata the PPV was less than 5%. Stellon and Kenwright studied all patients aged >50 years, where they found the PPV of iron deficiency anaemia for colorectal cancer to be greater than 5%.27

Urological. Three high-quality studies examined symptoms of urological cancer (Table 6). One of these studies excluded prostate cancer,27 one concerned undefined urological cancer,28 and the third only prostate cancer.27 Jones et al found PPVs in men aged ≥55 years and women aged ≥65 years to be greater than 5% for haematuria.27 Bruyninckx et al found macroscopic haematuria to have a PPV of over 5% in men aged ≥60 years and women aged ≥40 years.27 The study by Hamilton et al of prostate cancer only was a case-control design.27 They found a ‘malignant rectal

Figure 2. PPV of rectal bleeding for colorectal cancer in older adults. Age ranges for studies (years): ≥45 du Toit,6 Jones,10 Parker;7 ≥60 Ellis;8 ≥50 Wauters;10 Heintze;12 Hamilton (09); ≥40 Mant;6 Norrelund,12 Metcalf;4 Hamilton (05);22 50–75 Fijten.8 Figures for Hamilton 09 were estimated from original data sent by the authors.

Figure 3. PPV of rectal bleeding for colorectal cancer age 70–75 years and over. Age ranges for studies (years): ≥75 Du Toit,9 Jones,10 Parker;7 Hamilton;6 ≥70 Wauters.6 Figures for Hamilton 09 were estimated from original data sent by the authors.
examination’ to have a PPV of greater than 5% for prostate cancer. Some combinations of symptoms had a PPV above 5% for prostate cancer, but the authors did not calculate CIs for PPVs, because of small sample sizes.

**Lung.** One study identified haemoptysis as having a PPV of 5% or more for lung cancer, and another study resulted in a PPV for this symptom of less than 5% (Table 7). In the latter study, haemoptysis in combination with other symptoms, or recurrent haemoptysis, had high predictive values. However, the authors did not calculate CIs due to small sample sizes. Jones et al used only computer codes to identify the symptom, while Hamilton et al used codes and free text. Both studies were high quality.

**Oesophageal.** Two studies (one high and one low quality) were identified for oesophageal cancer (Table 8). No studies were identified that gave a PPV of less than 5% for the symptom of dysphagia.

**Breast.** Two studies (one high quality) identified a breast lump symptom, and one high-quality study a clinician-palpable breast lump, as having a PPV of 5% or more for breast cancer (Table 9). No studies were identified that gave a PPV of less than 5% for this symptom or sign. Two of the studies included all women, even though breast cancer in those aged <20 years is exceptionally rare.

**Gynaecological.** One high-quality retrospective study, using routinely recorded computer codes in the medical records, identified the first ever consultation for postmenopausal bleeding in women aged 75 to 84 years as having a PPV for relevant cancer of 5.4%...
Systematic Review

M Shapley, G Mansell, JL Jordan, KP Jordan

(95% CI = 4.7 to 7.1) (Table 10). No studies were identified that gave a PPV of less than 5% for this symptom in the age group 75 to 84 years.

Other cancers. No other studies were identified with PPVs of over 5% for symptoms, signs, or tests for other sites of cancer.

(95% CI = 4.7 to 7.1) (Table 10). No studies were identified that gave a PPV of less than 5% for this symptom in the age group 75 to 84 years.

F = female. Hb = haemoglobin (g dl⁻¹). M = male. MCV = mean corpuscular volume (fL). NA = not applicable. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value. RBC = red blood corpuscles.
Table 7. Studies giving a PPV of ≥5% in one or more strata of symptoms for lung cancer or a predictive value of <5% in an equivalent stratum.

<table>
<thead>
<tr>
<th>Study and NOQAS star rating</th>
<th>Type of data</th>
<th>Source</th>
<th>Follow-up</th>
<th>Age, years</th>
<th>Sex</th>
<th>Symptom</th>
<th>Age, years</th>
<th>Number with symptom</th>
<th>PPV % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones et al, 2007&lt;sup&gt;7&lt;/sup&gt; (9 stars)</td>
<td>Retrospective cohort</td>
<td>Routinely recorded computer codes</td>
<td>3 years</td>
<td>≥15</td>
<td>All</td>
<td>Haemoptysis</td>
<td>M</td>
<td>55–64</td>
<td>514</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65–74</td>
<td>552</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>75–84</td>
<td>393</td>
<td>17.1 (13.5 to 21.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>≥85</td>
<td>93</td>
<td>20.4 (12.8 to 30.1)</td>
</tr>
<tr>
<td>Hamilton et al, 2005&lt;sup&gt;5&lt;/sup&gt; (9 stars)</td>
<td>Case-control</td>
<td>Routinely recorded data including free text</td>
<td>Not applicable</td>
<td>≥40</td>
<td>All</td>
<td>Haemoptysis as a single symptom</td>
<td>M+F</td>
<td>≥40</td>
<td>Cases 50, controls 19</td>
</tr>
</tbody>
</table>

F = female. M = male. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value.

Table 8. Studies giving a PPV of ≥5% in one or more strata of dysphagia for oesophageal cancer.

<table>
<thead>
<tr>
<th>Study and NOQAS star rating</th>
<th>Type of data</th>
<th>Source</th>
<th>Follow-up</th>
<th>Age, years</th>
<th>Sex or sign</th>
<th>Symptom</th>
<th>Age, years</th>
<th>Number with symptom</th>
<th>PPV % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones et al, 2007&lt;sup&gt;7&lt;/sup&gt; (9 stars)</td>
<td>Retrospective cohort</td>
<td>Routinely recorded computer codes</td>
<td>3 years</td>
<td>≥15</td>
<td>M+F</td>
<td>Dysphagia</td>
<td>M</td>
<td>55–64</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65–74</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>75–84</td>
<td>476</td>
<td>7.1 (5.0 to 9.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>≥85</td>
<td>154</td>
<td>9.7 (5.6 to 15.6)</td>
</tr>
<tr>
<td>Esfandyari et al, 2002&lt;sup&gt;5&lt;/sup&gt; (4 stars)</td>
<td>Prospective cohort</td>
<td>Specifically registered</td>
<td>Not given</td>
<td>Not given</td>
<td>M+F</td>
<td>Dysphagia</td>
<td>M+F</td>
<td>Not given</td>
<td>100</td>
</tr>
</tbody>
</table>

F = female. M = male. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value.

Table 9. Studies giving a PPV of ≥5% in one or more strata of symptoms for breast cancers.

<table>
<thead>
<tr>
<th>Study and NOQAS star rating</th>
<th>Type of data</th>
<th>Source</th>
<th>Follow-up</th>
<th>Age, years</th>
<th>Sex or sign</th>
<th>Symptom or sign</th>
<th>Age, years</th>
<th>Number with symptom</th>
<th>PPV % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eberl et al, 2008&lt;sup&gt;8&lt;/sup&gt; (5 stars)</td>
<td>Retrospective cohort</td>
<td>Routinely recorded data</td>
<td>Not given</td>
<td>Not given</td>
<td>F</td>
<td>Breast lump/ mass symptom</td>
<td>F</td>
<td>Not given</td>
<td>741</td>
</tr>
<tr>
<td>Barton et al, 1999&lt;sup&gt;9&lt;/sup&gt; (8 stars)</td>
<td>Retrospective cohort</td>
<td>Routinely recorded data</td>
<td>1–11 years</td>
<td>40–69</td>
<td>F</td>
<td>Breast lump/ mass symptom</td>
<td>F</td>
<td>40–69</td>
<td>196</td>
</tr>
<tr>
<td>Bywaters 1977&lt;sup&gt;6&lt;/sup&gt; (6 stars)</td>
<td>Retrospective cohort</td>
<td>Routinely recorded data</td>
<td>Not given</td>
<td>All</td>
<td>F</td>
<td>Breast lump/ mass sign</td>
<td>F</td>
<td>All</td>
<td>57</td>
</tr>
</tbody>
</table>

F = female. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value.

Table 10. Studies giving a PPV of ≥5% in one or more strata of symptoms for gynaecological cancer.

<table>
<thead>
<tr>
<th>Study and NOQAS star rating</th>
<th>Type of data</th>
<th>Source</th>
<th>Follow-up</th>
<th>Age, years</th>
<th>Sex or sign</th>
<th>Symptom or sign</th>
<th>Age, years</th>
<th>Number with symptom</th>
<th>PPV % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parker et al, 2007&lt;sup&gt;7&lt;/sup&gt; (7 stars)</td>
<td>Retrospective cohort</td>
<td>Routinely recorded computer codes</td>
<td>2 years</td>
<td>≥40</td>
<td>F</td>
<td>First ever consultation with postmenopausal bleeding computer codes</td>
<td>Female</td>
<td>75–84</td>
<td>856</td>
</tr>
</tbody>
</table>

F = female. NOQAS = Newcastle-Ottawa Quality Assessment Scales. PPV = positive predictive value.
DISCUSSION

Summary of main findings

The study identified 21 papers that reported a PPV of 5% or more for a symptom, sign, or non-diagnostic test result alone or in combination in unselected primary care populations for cancer, and a further four studies reporting a PPV of less than 5% for the same symptom, sign, or non-diagnostic test result.

Strengths and limitations of the review

The definition of symptom differed between studies. In some, only the computer codes allocated to the consultation in the medical records were used, while others used the whole medical record including free text. The latter method will detect a greater number of patients with symptoms, as the computer code generally represents a summation of the consultation into a diagnosis or term representing patient management. This approach is encouraged in research practices (personal communication, Royal College of General Practitioners’ [RCGP] Weekly Returns Service, 2010), with a stated preference for GPs to record diagnostic computer codes rather than symptom codes. It is unusual for a GP to record multiple history or symptom codes for a single problem. The studies that calculate PPVs for computer codes thus give values for patients in whom a GP is unable to make a more specific diagnosis. These studies do not give PPVs for symptoms presented to GPs, unlike studies in which free text is also analysed. Both types of study overestimate the PPV as a consequence of under-recording by GPs, especially when the clinician has a low level of suspicion for cancer. The size of this effect is not known. The PPVs will only remain valid if GP recording and patient consultation behaviour do not change with time.

A significant source of bias in studies using clinicians to specifically register a patient is selection of those with more severe symptoms and those in whom the clinician has a greater suspicion of malignancy. Such studies are considered to have a low weight in the present review of unselected primary care populations due to this selection bias, and are likely to overestimate the PPV. In the meta-analysis of rectal bleeding, there is a large degree of heterogeneity and a trend for the PPV to be lower in more recent studies. This partially reflects this source of bias and a change in recruitment methods over time from the use of specifically registered patients to routinely recorded data. In studies using computer codes, the actual codes used were not reported and may differ significantly between studies even when the symptom appears to be a single entity (personal communications, W Hamilton and R Jones, 2009).

A recent systematic review and meta-analysis on studies that only involved the recruitment of cases of rectal bleeding by GPs commented on the difficulties of complete identification and follow-up of all incident cases of rectal bleeding in primary care and differences in age cut-off points between studies. The authors gave an estimated PPV of over 5% in those aged ≥60 years with rectal bleeding.

Several other methodological problems can overestimate the PPV in studies, including the use of cancer registries resulting in the inclusion of cancers not diagnosed through primary care (for example, screen detected, referral within hospital, and admission via accident units), cancers diagnosed as an association and not the cause of symptoms, and recall bias in case-control studies.

Underestimation of the PPV may occur with inadequate follow-up, and the length of follow-up varied between studies from no follow-up to 11 years. The study by Jones et al indicates that

![Table 11. Symptoms, signs and non-diagnostic tests in unselected primary care populations with a PPV of ≥5% for cancer for which there is robust evidence.](image-url)

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Cancer</th>
<th>Sex</th>
<th>Age, years</th>
<th>Evidence level for PPV of 5% or more in other cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectal bleeding computer code or new onset rectal bleeding</td>
<td>Colorectal</td>
<td>M+F</td>
<td>≥75</td>
<td>None aged &lt;60 years, aged 60–74 years equivocal</td>
</tr>
<tr>
<td>Iron deficiency anaemia Hb &lt;12 g/dl</td>
<td>Colorectal</td>
<td>M</td>
<td>≥60</td>
<td>Moderate evidence for gastrointestinal malignancy in men aged &lt;12 g/dl aged ≥60 years and women with Hb ≤11 g/dl aged &gt;50 years</td>
</tr>
<tr>
<td>Iron deficiency anaemia Hb &lt;11 g/dl</td>
<td>Colorectal</td>
<td>F</td>
<td>≥70</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Iron deficiency anaemia Hb &lt;9 g/dl</td>
<td>Colorectal</td>
<td>F</td>
<td>≥60</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Haematuria</td>
<td>Urological</td>
<td>M+F</td>
<td>≥60</td>
<td>None aged &lt;40 years, aged 40–60 years equivocal</td>
</tr>
<tr>
<td>Rectal examination malignant</td>
<td>Prostate</td>
<td>M</td>
<td>≥40</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Haemoptysis computer code</td>
<td>Lung</td>
<td>M</td>
<td>≥55</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Haemoptysis computer code</td>
<td>Lung</td>
<td>F</td>
<td>≥65</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Dysphagia computer code</td>
<td>Oesophagus</td>
<td>M</td>
<td>≥55</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Breast lump or mass</td>
<td>Breast</td>
<td>F</td>
<td>≥20</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
<tr>
<td>Postmenopausal bleeding computer code</td>
<td>Gynaecological</td>
<td>F</td>
<td>75–84</td>
<td>No other evidence for PPV of 5% or more in other cohorts</td>
</tr>
</tbody>
</table>

F = female. Hb = haemoglobin. M = male. PPV = positive predictive value.
Some cancers are diagnosed 3 years after initial presentation.12

Other potential sources of bias in this review were that it only included English language papers, and that some papers in more specialist journals could not be located. This potential bias is unlikely to affect the results of the review, as, on closer examination of the abstracts, they are unlikely to have matched the study eligibility criteria.

The studies identified that calculated PPVs for combinations of symptoms, signs, or non-diagnostic test results and risk of cancer, involved populations that were too small and/or methods of recruitment that were subject to too high a risk of bias to draw firm conclusions. Three recent systematic reviews on the diagnostic utility of combinations of symptoms, signs, and test results for colorectal cancer showed that evidence in primary care is lacking, and when present only shows modest diagnostic value.28-30

A potential criticism of this review is that it only included symptoms, signs, and tests with a point estimate PPV of 5% or more. It is likely that it excluded good-quality studies with a point estimate of less than 5%, but in which the 95% CIs included 5%. It is possible that future research may identify additional symptoms, signs, and tests with a point estimate of 5% or more.

The level of risk an individual is prepared to take is specific to that person at that moment, and is dependent on many factors. Twenty-six additional studies were identified with PPVs of other symptoms, signs, and non-diagnostic test results for cancer of less than 5%. A lower level of PPV than that used in this study will identify more symptoms, signs, and non-diagnostic test results for cancer, and additional age and sex groups in unselected primary care populations. Selecting a lower level of risk may diagnose cancer earlier in presenting populations but with the disadvantage that more individuals will undergo negative investigation with physical, psychological, and economic cost from the tests. A maximum level of risk for a population of 5% was chosen for this review, based on stated levels of risk in national guidelines concerning a comparable disease and treatment. The lower the level of risk, the greater the disagreement as to whether or not inaction can be justified.

When communicating with patients, it is preferable to use absolute measures of risk (number needed to investigate) rather than relative measures of risk.41 In assessing a patient’s level of risk using predictive values, it is important to individualise the risk, as the PPV of the population will be modified by the individual’s risk factors for the illness. There has been no debate with the public regarding what level of risk is acceptable.41

Implications for future research and clinical practice

This study found evidence for only nine symptoms, signs, and non-diagnostic test results that have a high predictive value for cancer in patients presenting to GPs during routine consultations. A number of studies were subject to methodological and sample size difficulties, and for one symptom (change in bowel habit and colorectal cancer), high-quality studies using different methodology produced contradictory results.

On the basis of good quality studies (one or more studies on the quality-assessment tool at a level of 6 stars or more and no similar quality contradictory studies), broad consistency with published guidelines,4 and the unlikelihood that further research will change the conclusion of studies performed on large databases, it can be concluded that robust evidence exists for a PPV of 5% or more in specific age/sex groups as detailed in Table 11.

The low number of symptoms, signs, and non-diagnostic test results with a PPV of 5% or more was not unexpected, as the difficulties in diagnosing cancer in primary care have been commented on before.3 While these symptoms, signs, and non-diagnostic test results occur in the minority of patients with cancer, when they do occur, exclusion of cancer is obligatory unless exceptional circumstances exist.

There is some evidence that GPs do not always act upon them.34,42,43 A previous effort to trigger GP action concerning one of the tests was unsuccessful,44 but this may have been due to the form of trigger used. The identified symptoms, signs, and non-diagnostic test results are broadly consistent with NICE recommendations for referral for suspected cancer. However, NICE generally uses a lower threshold for referral for additional symptoms and tests, and in the case of haemoptysis, even when the computer code is associated with a PPV of 5% or more.

The authors recommend research and development of general practice computer systems to produce effective warning flags when the symptoms, signs, or test results with a PPV of 5% or more from unselected primary care populations are entered for patients within the specified sex and age groups. The management of these patients should be audited and appraised. There is a need to standardise terms and analyses, and for further research on combinations of symptoms, signs, and non-diagnostic test results. There should be more open debate on the level of PPV that triggers a recommendation for referral by a GP.

Funding body

NHS North Staffordshire; NHS Stoke; NHS Executive West Midlands R&D Office; North Staffordshire and Cheshire R&D Consortium.

Ethical approval

Not applicable.
Competing interests

The authors have stated that there are none.

Acknowledgements

We wish to thank Drs Judith Bell and Giri Rajaratnam, Professor Peter Croft and Brian Dudley for commenting on the study design and draft paper.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: http://www.rcgp.org.uk/bjgp-discuss

REFERENCES