Bladder cancer is common worldwide. Worldwide, there are nearly 400 000 new cases diagnosed annually, leading to over 150 000 deaths.\(^1\) It is strongly associated with cigarette smoking, and is approximately 2.5 times more common in males.\(^2\) The incidence rises with age, with the average age at diagnosis being 71 years.\(^2\) At the time of diagnosis in the UK, approximately 75–85% patients have a non-invasive tumour, with a recurrence rate of 31–78%, yet with high 5-year survival of 80–90%. Around 30% are multifocal. In contrast, once the tumour has invaded muscle, survival is below 50%.\(^3\) Emergency presentations occur predominately with advanced disease and have a higher mortality.\(^3\)

More timely diagnosis of bladder cancer may improve outcomes, either by a favourable stage shift or by avoiding emergency presentations. There is some evidence to suggest that better survival is seen in those with shorter times to diagnosis.\(^4,5\) This may be particularly relevant to the UK and other countries with a gatekeeper system, where access to specialist care requires referral from primary care clinicians. These countries have worse cancer outcomes.\(^4\) There were an estimated 752 additional bladder cancer deaths in the UK for 1995–1999 when compared with the European average,\(^7\) although some improvement has been seen in selected cancers recently.\(^8\)

There is no accepted screening test for bladder cancer, so diagnosis currently requires presentation with symptoms. This is usually to primary care.\(^9\) However, most diagnostic studies have originated in secondary care. As this population has already been selected for investigation, such results cannot illuminate the selection process. As well as selection and possible recall biases, studies of the referred population inevitably examine patients later in the progress of their disease. The few primary care studies have only examined haematuria. In the first, all 363 patients over 18 with haematuria (either visible or non-visible) presenting to primary care in Hull, UK, were studied.\(^10\) Cancers were found in 36, with 28 (77% of the cohort) being bladder tumours. Only three of the 186 with non-visible haematuria transpired to have urological cancer (1.6%), in contrast with 32 of the 172 (19%) with visible haematuria (the minor numerical inconsistency being the authors’). Cancer was more common in patients with a history of urinary tract infection, and slightly so in those with obstructive urinary symptoms. A second study in Belgian primary care identified all patients with visible haematuria, and calculated a positive predictive value (PPV) of 150 000 deaths worldwide. No screening is available, so diagnosis depends on investigations of symptoms. Of these, only visible haematuria has been studied in primary care.

Visceral haematuria is the commonest and most powerful predictor of bladder cancer in primary care, and warrants investigation. Most other previously reported features of bladder cancer were associated with the disease, but with low predictive values. There is a need for improved diagnostic methods, for those patients with bladder cancer presents without visible haematuria.
How this fits in

The UK has no screening for bladder cancer so diagnosis in primary care relies on symptomatic presentation. Recognising the early symptoms of bladder cancer could improve the UK’s poor mortality outcomes. Existing research has identified haematuria as a significant indicator of risk. No primary care study has looked at identifying multiple features of bladder cancer, including investigations. This study aimed to identify and quantify all significant features in primary care patients. The sample size of nearly 5000 cases identified eight features that were significantly associated with bladder cancer. Results were quantified using a risk assessment tool (RAT) designed to aid GP’s referral decision making. Visible haematuria presented the highest risk as a single feature with a positive predictive value (PPV) of 2.6%, therefore warranting referral. Risks increased with multiple presentation and when combined with abnormal investigations. These results are being fed back in to the re-write of the NICE guidelines.

for bladder cancer of 8.3%. Additional symptoms did not alter this figure. In a third study, PPVs for bladder cancer of 8.0% for males and 3.7% in females were calculated for visible haematuria in a large cohort aged 16–100 years. Several other features of bladder cancer have been described in the secondary care literature, including lower urinary tract symptoms, abdominal pain, and masses.

The aim of this study was to identify the symptoms of bladder cancer in primary care, and to quantify them, both singly and in pairs. This should allow improved selection of patients for investigation, reducing the human and economic cost of this disease.

METHOD

This was a matched case-control study using patients in the UK’s General Practice Research Database (GPRD). The GPRD contains approximately 5 million primary care patient records from general practices around the UK, for good geographical representativeness. All primary care consultations are recorded including symptoms, diagnoses, and prescriptions. Investigation results are also documented. Only general practices adhering to stringent data quality recording levels are included in the GPRD.

Cases and controls

Cases had a first record of bladder cancer (using one of 20 medical codes for bladder cancer used in the GPRD — available from authors) between January 2000 and December 2009 inclusive. They were aged ≥60 years, and had a minimum of 1 year of data before diagnosis. The first instance of a bladder cancer code was assigned the date of diagnosis: the index date. The index date for controls was the index date for their matched case. Up to five controls were matched on sex, general practice, and to 1 year of age of the case. The GPRD only provides the year of birth (to protect anonymisation), so an arbitrary birth date of 1 July was assigned to everyone. Exclusion criteria were: metastatic cancer of the bladder from a non-bladder primary, diagnosis before 2000, or no consultations in the year before diagnosis.

Selection of investigations and symptom variables

A list of symptoms, signs, and investigations (called ‘features’ from now on) potentially associated with bladder cancer was compiled from the literature search, augmented by viewing material from bladder cancer support organisations and online chat rooms. Internet search terms included ‘bladder cancer’, ‘bladder symptoms’, and ‘early signs/indications’. Visible and non-visible haematuria were studied separately. Only codes specifying the word ‘microscopic’ were assigned to the latter group, so generic codes such as the single word ‘haematuria’ were assumed to be visible haematuria. For each feature a list of relevant medical codes was assembled from the GPRD’s master list of over 100 000 codes. Occurrences of these were identified in the year before the index date. Repeated consultations for the same complaint were also identified. Also identified were all codes for fractures as a test for any recording bias between cases and controls (making the assumption that the fracture rate would be approximately equal). Variables were retained only if they occurred in at least 5% of either cases or controls [this was always cases]. Investigation results were deemed to be abnormal if they fell outside their local laboratory’s normal range for analysis. Patients with a normal laboratory result were grouped with those who had not been tested.

Analysis and statistical methods

This followed the methods used in several previous studies. The main analytical method was conditional logistic regression. Analysis was performed in three stages. All variables associated with cancer with a
A *P*-value of ≤0.1 in univariable analysis were retained. In the second stage, variables were grouped for multivariable analysis, collecting together variables that were similar, such as visible and non-visible haematuria, using a *P*-value threshold of ≤0.05. The final stage of multivariable analysis used all variables surviving the previous stages, and used a *P*-value threshold of 0.01. All excluded variables were checked against the final model.

Positive predictive values (PPVs) were estimated for features shown to be independently associated with cancer in the multivariable analysis. This was repeated for pairs of symptoms and for second attendances with the same symptom. PPVs were estimated using Bayes’ theorem, whereby the prior odds x likelihood ratio (LR) = posterior odds. The prior odds used were the age-specific national incidence of bladder cancer for 2008, expressed as odds. To enable a calculation of PPVs for the consulting population, the proportion of the control population who had not consulted in the year before diagnosis was estimated. Of 23,804 eligible controls, 2,086 (8.8%) had not consulted; so PPVs were divided by 0.912 to give the figure for the consulting population (Figure 1).

As the number of cases in the GPRD were fixed, a power calculation was performed rather than a sample size calculation. Thus, 5,000 cases and 20,000 controls (the estimates initially provided by the GPRD) provided >99.9% power (5% two-sided alpha) to detect a change in a rare variable from 1% in cases and 2% of controls. For a commoner variable, the study had >89% power to detect a change in prevalence of 20% in cases to 18% in controls. Data analysis was conducted using Stata software (version 11).

## RESULTS

The GPRD provided 29,033 patients (4,935 cases; 24,098 controls). Application of the exclusion criteria is shown in Figure 1, leading to a final number of 26,633 (4,915 cases; 21,718 controls).

Patient demographic and consultation information is given in Table 1. Cases consulted significantly more frequently than controls in the year before diagnosis (*P* ≥ 0.001; rank sum test).

### Clinical features

Forty-three symptoms and 104 abnormal test results were considered initially. Six symptom and seven abnormal test variables were present in at least 5% of cases. Only 2.6% of cases in the study had a record of non-visible haematuria. Features associated with bladder cancer in univariable analysis are shown in Table 2. Raised inflammatory markers were defined as being above the local laboratory’s quoted normal range for erythrocyte sedimentation rate or C-reactive protein. Table 2 also shows the multivariable analysis. All variables retained an independent association with cancer, with a *P*-value <0.001 in the final multivariable model. No interaction terms were found.

From a total of 4,915 cases, 3,621 (74%) had at least one of the features from the final model in Table 2 recorded in their notes. The proportion of patients with a recorded fracture did not differ between cases (n = 71, 1.45%) and controls (n = 316, 1.46%).

### Positive predictive values

Figure 2 shows the PPVs for individual, combined, and repeat features, for patients aged ≥60 years. The LRIs were largely similar between the two age groups (40–59 years and ≥60 years) except for visible haematuria, which had a stronger association with cancer in younger patients (LR 1.89) compared with older patients (LR 3.8). The incidence of bladder cancer is approximately tenfold...
lower in patients aged 40–59 than in those aged ≥60 years, so PPVs in the younger age group are also approximately tenfold lower for symptoms other than visible haematuria (data not shown). The PPV of visible haematuria in the 40–59 year age group was 1.4% [95% confidence interval (CI)= 0.5 to 4.4).

**DISCUSSION**

This is the first study from primary care to identify all the clinical features of patients with bladder cancer in that setting. It was possible to quantify these features using the most useful metric for clinicians: the PPV for a consulting patient. Four symptoms were identified that were both common and significantly associated with bladder cancer, from an original list gathered from several sources as well as conventional publications. Also confirmed was an association for urinary tract infection and three standard investigations. For a patient aged ≥60 years describing haematuria to his or her GP, the risk of cancer was 2.6%. Risks increased with multiple and repeated symptoms. In the absence of visible haematuria or dysuria, however, risks were consistently low.

**Table 1. Patient demographics and consultation rates in the year before diagnosis**

<table>
<thead>
<tr>
<th></th>
<th>Case</th>
<th></th>
<th>Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Total</td>
<td>Male</td>
</tr>
<tr>
<td>Median age at diagnosis, years (IQR)</td>
<td>73 (65–80)</td>
<td>75 (67–82)</td>
<td>74 (66–80)</td>
<td>73 (66–79)</td>
</tr>
<tr>
<td>Median number of consultations (IQR)</td>
<td>14 (9–22)</td>
<td>15 (10–23)</td>
<td>15 (9–22)</td>
<td>8 (4–15)</td>
</tr>
</tbody>
</table>

**Table 2. Clinical features of bladder cancer (all ages)**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cases, n (%)</th>
<th>Controls, n (%)</th>
<th>LR(^a) (95% CI)</th>
<th>OR(^b) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible haematuria</td>
<td>2595 (53)</td>
<td>196 (1)</td>
<td>59.5 (51 to 67)</td>
<td>34.2 (29 to 41)</td>
</tr>
<tr>
<td>Dysuria</td>
<td>444 (9)</td>
<td>209 (1)</td>
<td>9.4 (8.0 to 11)</td>
<td>4.1 (3.4 to 5.0)</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>358 (7)</td>
<td>787 (4)</td>
<td>2.0 (1.8 to 2.3)</td>
<td>2.0 (1.6 to 2.4)</td>
</tr>
<tr>
<td>Constipation</td>
<td>286 (6)</td>
<td>708 (3)</td>
<td>1.8 (1.6 to 2.0)</td>
<td>1.5 (1.2 to 1.9)</td>
</tr>
<tr>
<td>Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>835 (17)</td>
<td>705 (3)</td>
<td>5.2 (4.8 to 5.8)</td>
<td>2.2 (2.0 to 2.5)</td>
</tr>
<tr>
<td>Investigations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raised creatinine</td>
<td>660 (13)</td>
<td>1668 (8)</td>
<td>1.8 (1.6 to 1.9)</td>
<td>1.3 (1.2 to 1.4)</td>
</tr>
<tr>
<td>Raised inflammatory markers</td>
<td>293 (6)</td>
<td>717 (3)</td>
<td>1.8 (1.6 to 2.1)</td>
<td>1.5 (1.2 to 1.9)</td>
</tr>
<tr>
<td>Raised white blood cell count</td>
<td>250 (5)</td>
<td>401 (2)</td>
<td>2.8 (2.4 to 3.2)</td>
<td>2.1 (1.6 to 2.8)</td>
</tr>
</tbody>
</table>

\(^a\) the univariate likelihood ratio, showing the likelihood of having a specific feature in a patient with bladder cancer, compared with the likelihood of having it in a patient without cancer. \(^b\) in multivariate conditional logistic regression, containing all eight variables. LR = likelihood ratio. OR odds ratio.

**Strengths and limitations**

The main strength of this study is its primary care setting. This is where the clinical problem exists: the decision of whom to select for investigation for possible bladder cancer. A further strength is that the symptom and investigation variables were collected before the diagnosis was established. Such a design eliminates recall bias (which can affect studies reporting retrospectively collected data), although it may add recording bias (see below). The study was large and generalisable, as GPRD data are representative of patients across the UK.

The main limitation of using electronic records as the data source is reliance upon the quality of the doctors’ recording. Patient records have been widely used in similar cancer studies before, including GPRD data.12,14–16,18 Symptoms may be unvoiced or unrecorded, but this only matters if one of these is more common in either of the two groups, cases or controls. There was reason to suspect this, and the test — the frequency of recording of fractures in cases and controls — supported this. If under-recording is equally prevalent between cases and controls, the likelihood ratios and PPVs will still be correct, although the 74% of patients with at least one feature of bladder cancer will be an underestimate. Indeed, PPVs from electronic primary care studies like this one largely match those derived from the few truly prospective studies.20,21 The matched design meant that controls were from the same practice as cases, which should reduce recording bias. Matching makes it impossible to study the matched variables directly, although the study was large enough for sub-analyses by sex and age-group to be performed, sidestepping that potential weakness. Furthermore, laboratory data have been transmitted automatically to practices since around 2000, so there should be no recording bias (which can affect studies reporting retrospectively collected data), although it may add recording bias present for these variables, which is why that date was chosen for the start of the study.

Only symptoms previously reported with bladder cancer were studied, although including patient groups and online chat rooms in a literature search made it unlikely that any salient symptoms were omitted. The investigations were even more inclusive: all primary care tests entered the analysis, although few were found to be significant. Verification bias occurs when testing is more commonly performed in one group (generally cases) than the other, so positive results are more common as a result of more testing (as well as any true
association). This could have been possible for non-visible haematuria, whereby a doctor may be more likely to perform urinalysis in a patient with bladder cancer. This variable was only present in 2.6% of cases, so was omitted from the main analysis.

Finally, the methods for calculating PPVs are innovative, although are now well-established.\textsuperscript{14} It is impossible to estimate PPVs using only the information within a case-control study (as the proportion of cases — the prior odds — is higher than in the population as a whole). In cancer studies, excellent national incidence statistics are available as a proxy for the prior odds. Indeed, these national figures may actually be superior to the prior odds derived from any cohort study in primary care, as the number of cancers in such a study would perform be very small.

**Comparison with previous literature**

The only symptom of bladder cancer previously reported from primary care is visible haematuria. A study of patients in a UK open-access clinic for both visible and non-visible haematuria identified several additional factors increasing the likelihood that a urological cancer was present.\textsuperscript{10} These included age and sex, several lower urinary tract symptoms, and a history of previous urinary tract infection. No PPVs were estimated. The Belgian study estimated a PPV of 10.3% for urological cancer with visible haematuria presented to primary care; 69% of these were bladder cancers, giving an approximate 7% PPV for bladder cancer. The figure of 1.4% for ages 40–59 years was based on very few controls reporting visible haematuria to their doctors and so has very wide CIs. However, those aged ≥60 years had a PPV of 2.6% (95% CI = 2.2 to 3.2) as numbers were much higher. Overall, the Belgian study had considerably fewer cases — only 87 — and thus had wide CIs for all their risk estimates. Their cases of cancer were identified from a cancer registry quite separate from the collection method for those reporting haematuria. Further, Belgian patients may access urologists directly, so those patients reporting haematuria to primary care may be a self-selected group.\textsuperscript{22} Few clinicians, or their patients, would disagree with investigation for cancer at a risk of 2.6%, and many would choose investigation at a lower value. Therefore, this study supports investigation for possible bladder cancer in all the over-40s with visible haematuria. This is the view of NICE guidance and a recent review, but without primary care data to substantiate them.\textsuperscript{23,24} A 2003 systematic review of the economics of bladder cancer identified six papers examining the cost-effectiveness of investigation for bladder cancer.\textsuperscript{25} In a population with a prevalence of 1% there was an estimated cost of $170 000 per case identified; this figure fell to $12 000 per case if the prevalence were 15%.\textsuperscript{26} No recent evaluation has been reported, although it is likely the “threshold” level for investigation should be above 1%, on health-economic grounds at least.

A link between urinary tract infection and bladder cancer has been reported before: either with infection predisposing to cancer, or infection being a presenting feature of a cancer.\textsuperscript{27} This study only addressed the association.)

**Figure 2. Positive predictive values for bladder cancer in patients aged ≥60 years, for individual, paired, and repeated features.**

![Figure 2. Positive predictive values for bladder cancer in patients aged ≥60 years, for individual, paired, and repeated features.](image-url)

The PPV is shown on the first line of each cell with the 95% CIs shown underneath. Where fewer than 10 cases or controls had the combined features, CIs were omitted. The yellow-shaded cells indicate a PPV of at least 1.0%, orange 2.0%, and red 5.0% or over. The cells showing the same feature vertically and horizontally represent a second attendance with the same investigation, so are omitted for test results.
latter aspect, showing a moderate increase in the risk of cancer when urinary infection was recorded by a GP. Also found were independent associations for dysuria as a symptom, and for raised inflammatory markers or leukocytosis, all of which may be present with urinary infection. Indeed, this distinction between urine infection and dysuria may reflect the data source, in that different doctors may record similar clinical scenarios differently. Nonetheless, it is clear that such symptoms are relevant in diagnosis of possible bladder cancer. A classic cause of delay in bladder cancer is when a patient has repeated urine infections, each successfully treated, but without bladder cancer being considered. The study results show that urological anecdote has some basis in truth. Two other symptoms were roughly twice as common in cases as controls: constipation and abdominal pain. However, as isolated symptoms — or even together — their predictive values were very small, so it is unlikely they will provide much help in selection of patients for investigation.

Decreased renal function has previously been reported with bladder cancer. It cannot be determined if the patients had obstructive nephropathy, to account for the higher percentage of cases in this study with raised creatinine.

Implications for research and practice

The findings support investigation of all patients aged >40 years with visible haematuria. It is highly likely that such a recommendation would be supported on economic as well as clinical grounds. The only other clinical scenario with a risk of cancer above 1% was repeated attendances with dysuria in patients >60 years. Figure 2 adds to the other cancer ‘Risk Assessment Tools’ (RATs) that are being introduced into UK primary care. Currently, only colorectal and lung RATs are being disseminated. If they are found to be helpful, bladder could easily be added. This may best be done allied with prostate, already published in the BJGP, and kidney, which is currently being studied, into a single urological RAT.

Funding

This study was funded by the NIHR Programme Grants for Applied Research funding scheme, RP-PG-0608-10045. William T Hamilton was part-funded by a NIHR post-doctoral fellowship.

Ethical approval

Independent Scientific Advisory Committee: protocol 09-110.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors would like to acknowledge the contribution to the research presented in this paper made by the Discovery Programme Steering Committee comprising: Roger Jones (chair); Jonathan Banks; Alison Clutterbuck; Jon Emery; Joanne Hartland; Sandra Hollinghurst; Maire Justice; Jenny Knowles; Helen Morris; Tim Peters; Greg Rubin.

Discuss this article

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

13. Lawrenson R, Logie J, Marks C. What if cancer survival in Britain were the same as in Europe: how many deaths are avoidable? Br J Cancer 2009; 101(52): S115–124.