LUNG CANCER

What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study

W Hamilton, T J Peters, A Round, D Sharp

Thorax 2005;**60**:1059–1065. doi: 10.1136/thx.2005.045880

See end of article for authors' affiliations

Correspondence to: Dr W Hamilton, Academic Unit of Primary Health Care, Department of Community Based Medicine, University of Bristol, Bristol BS8 1AU, UK; w.hamilton@bristol.ac.uk

Received 26 April 2005
Accepted 20 September 2005
Published Online First 14 October 2005

Background: Over 38 000 new cases of lung cancer occur each year in the UK. Most are diagnosed after initial presentation to primary care, but the relative importance of the various clinical features is largely unknown.
Methods: A population based case-control study was undertaken in all 21 general practices in Exeter, Devon, UK (population 128 700). 247 primary lung cancers were studied in subjects aged over 40 years diagnosed between 1998 and 2002 and 1235 controls matched by age, sex and general practice. The entire primary care record for 2 years before diagnosis was coded using the International Classification of Primary Care-2. Univariable and multivariable conditional logistic regression analyses were used to identify and quantify clinical features independently associated with lung cancer. The main outcome measures were odds ratios and positive predictive values for these variables.
Results: Seven symptoms (haemoptysis, loss of weight, loss of appetite, dyspnoea, thoracic pain, fatigue and cough), one physical sign (finger clubbing), and two abnormal investigation results (thrombocytosis and abnormal spirometry) were associated with lung cancer in multivariable analyses, as was cigarette smoking. After excluding variables reported in the final 180 days before diagnosis, haemoptysis, dyspnoea and abnormal spirometry remained independently associated with cancer.
Conclusions: This study provides an evidence base for selection of patients for investigation of possible lung cancer, both for clinicians and for developers of guidelines.

Lung cancer is the most common cause of death from cancer in the industrialised world. Over 38 000 new cases of lung cancer occur each year in the UK,[1] and over 170 000 in the USA.[2] Mortality is related to the stage at diagnosis, with the best prognosis in early stage cancers. Earlier diagnosis of lung cancer may be beneficial in allowing some patients to have curative surgery and others with inoperable disease to have less extensive treatment. One possible route to earlier diagnosis is screening, although trials of screening using chest radiography have yielded disappointing results.[3] A large prospective trial comparing low dose spiral computed tomographic (CT) scanning with chest radiography in current or former smokers is due to report interim results shortly.[4] In the absence of screening, the main prospect for earlier diagnosis is prompt recognition of symptomatic cancer.[5] This will usually be in primary care but may occur in any healthcare setting.[6]

Several symptoms of lung cancer have been described in secondary care series of patients, but very little research has been reported on unselected populations such as primary care. One case series of 40 lung cancer patients in primary care found that 33% of cases had reported cough, 18% dyspnoea, 15% chest pain, and weight loss or fatigue each in 10%.[7] A recent UK series of 22 patients suggested much higher figures with 60–70% reporting symptoms of cough, fatigue, loss of appetite, chest pain, or dyspnoea.[8] Haemoptysis occurred in 41%. However, all the symptoms of cancer can also be found—and are much more common—in benign conditions. Approximately 5% of all primary care consultations are for cough[9] and 1.5% of the population consult for fatigue each year.[10] The incidence of haemoptysis (a cardinal symptom of lung cancer) has not been reported from primary care. This applies to the general population consulting their doctor as well as for those with lung cancer.[11] Furthermore, no studies in primary or secondary care have calculated predictive values for any of the symptoms of lung cancer.

The initial primary care investigation for a patient with possible lung cancer is chest radiography. However, this may occasionally fail to show the tumour. If suspicion of cancer remains, referral for other tests such as CT scanning or bronchoscopy may be required.[12] Selection of patients for investigation should ideally be based on knowledge of the risk posed by a particular symptom. It is not possible to use figures derived from secondary care to guide clinicians in primary care (or any other setting where unselected patients are managed) as the sensitivities, specificities, and predictive values for each symptom differ between these settings.[13 14] Despite the absence of relevant research, referral guidelines for suspected cancer have been established, advising on symptoms that should prompt consideration of a chest radiograph or referral to a respiratory physician.[5 15]

In the UK almost all the population receive primary care from National Health Service general (family) practitioners. These doctors maintain records of all primary and secondary care consultations. The records are of high quality and include the symptoms that patients have reported, as well as examination and investigation findings.[16 17] We designed a population based case-control study using these records with two main aims: (1) to identify the prediagnostic features of lung cancer; and (2) to calculate the positive predictive value of symptoms, physical signs, and abnormal test results for lung cancer in an unselected population.

METHODS

Subjects

Eligible cases were residents of Exeter, Devon, UK aged 40 years or over who had a primary lung cancer diagnosed

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.

during 1998–2002 inclusive. The total population of Exeter was 128 700 in mid 2000, of whom 44 561 were aged 40–69 years and 15 549 aged 70 years or above. Cases were identified from the cancer registry at the Royal Devon and Exeter Hospital which contributes to the South West Cancer Intelligence Unit. The percentage of lung cancer cases identified solely from death certification (so missed from registration) in Exeter primary care trust in 2002 was 2.9%; to this must be added those diagnosed and treated entirely at other hospitals which will have been a very small number. We sought to identify such additional cases by computerised searches at all 21 general practices in the city. Histological records were used to confirm the cancer, and those without positive histological results were accepted only if the records contained a specialist diagnosis of lung cancer based on strong clinical evidence. The date of diagnosis was taken as the date of positive histology or as that given by the specialist in those without histological proof.

Five controls were matched to each case using three criteria: sex, age, and general practice. Where more than five controls were available, the five were selected using computerised random numbers. Controls were eligible if they were alive at the time of diagnosis of their case; this did not preclude them from being dead at the time of study. Exclusion criteria for both cases and controls were (1) general practice record unobtainable; (2) no entry in the records in the 2 years before diagnosis; (3) the subject had a previous lung cancer; or (4) they lived outside Exeter at the time of diagnosis. Ineligible controls were replaced by randomly selected reserve controls. If an ineligible control was dead at the time of study they were replaced by a reserve control also known to be dead.

The study was approved by the North and East Devon research ethics committee.

Collection and coding of medical data

Anonymised photocopies of the full general practice records, both written and computerised, for 2 years before the date of diagnosis of each case were made. The records of dead patients were retrieved from storage by the local Health Authority. Four research assistants, blind to case/control status, coded all symptoms, physical signs and investigation results in the records using the International Classification of Primary Care-2.[18] This is the most symptom based of the common coding systems.[19] A small number of additional codes were created to incorporate all possible prediagnostic features. The same researcher coded both cases and controls within each general practice so that any inter-observer variation in coding style would affect both cases and their matched controls equally.

Past medical history and, where known, smoking and alcohol records were coded by a separate coder. Smoking records were accepted up to 5 years before diagnosis and subjects were categorised as non-smokers, ex-smokers, or current smokers. Chest radiographic results were collected but were not used in the main analyses as the doctor's decision to request a chest radiograph could imply that lung cancer was being considered as a possibility. Even if lung cancer was unsuspected, the chest radiograph would usually reveal it. Furthermore, by excluding chest radiographic results from the main analysis, the results can be used as a guide for when a chest radiograph should be considered.

Analysis of data

Identification of independent associations with cancer

Only variables occurring in at least 2.5% of either cases or controls were studied. Differences between cases and controls were analysed using conditional logistic regression. Variables associated with cancer in univariable regressions, using a p value of $\leqslant 0.1$, were entered in the multivariable regression analyses. These variables were placed in 10 clinical groups each containing 7–19 variables with a common theme such as pain, infection, or airways irritation. Each group was analysed by multivariable conditional logistic regression. Those variables remaining significantly associated with lung cancer after the first stage of analysis were regrouped and further modelling performed. All discarded variables were then checked against the final model. Sixteen clinically plausible interactions were tested in the final model. Analyses were repeated excluding data from the last 180 days of the 730 day period studied.

Calculation of positive predictive values

This was possible because we had identified almost all cases occurring in the population. Positive predictive values (PPVs) for individual variables and for pairs of variables were calculated from the likelihood ratio and the observed incidence of cancer during the study.[20] As all cases had consulted in primary care but 7.4% of initially selected controls had not, PPVs were divided by 0.926 to give the PPV for the consulting population. Confidence intervals (CIs) for these were calculated using Markov Chain Monte Carlo methods in WinBugs.[21] Stratified analyses by age (40–69 and 70+ years) were performed for individual features, but these were not performed if any cell in the 2×2 table was below 10.

Sample size calculations

Sample size calculations using a target of 250 cases gave 88% power to identify a change in a rare variable from 5% in controls to 11% in cases, and 85% power to identify a change in a common variable from 30% in controls to 40% in cases, both with a two sided 5% alpha. Analyses were performed using Stata Version 8.[22]

RESULTS

Cases and controls

A total of 299 cases were identified from the combined cancer registry (n = 296) and practice searches (n = 3) combined. Thirty nine cases were ineligible: two had previous lung cancer; 28 were unconfirmed or atypical cancers (22 of these were pleural mesotheliomas); seven were metastatic cancers from a non-lung primary; and two resided outside Exeter at the time of diagnosis. Of the 260 eligible cases, 13 could not be studied as the notes were unobtainable (five had left Exeter, eight had died and the notes could not be traced). This left 247 cases for whom 1235 matched controls were

Table 1 Characteristics of lung cancer cases and matched controls

Characteristic	Cases n = 247	Controls n = 1235
Age at diagnosis*		
<60	35 (14)	178 (14)
60–69	60 (24)	310 (25)
70–79	118 (48)	575 (47)
80+	34 (14)	174 (14)
Male sex*	170 (69)	850 (69)
Number of consultations per patient†		
In the 2 years before diagnosis	16 (9–24)	10 (5–17)
Excluding the last 180 days before diagnosis	9 (3–14)	8 (4–13)
Number of ICPC codes per patient†		
In the 2 years before diagnosis	34 (21–54)	20 (10–34)
Excluding the last 180 days before diagnosis	15 (6–31)	14 (7–26)

*Values shown as number (%).
†Values shown as median (interquartile range).

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.

Table 2 Frequency of selected variables in cases and controls

Characteristic	Cases (n = 247)*	Controls (n = 1235)*	Likelihood ratio	95% CI
Symptoms				
Haemoptysis	50 (20)	19 (1.5)	13	7.9 to 22
Loss of weight	67 (27)	54 (4.4)	6.2	4.5 to 8.6
Loss of appetite	47 (19)	49 (4.0)	4.8	3.3 to 7.0
Dyspnoea	139 (56)	192 (16)	3.6	3.1 to 4.3
Chest or rib pain	100 (42)	150 (9.2)	3.3	2.7 to 4.1
Fatigue	87 (35)	186 (15)	2.3	1.9 to 2.9
Consulting with cough				
First attendance	160 (65)	364 (29)	2.2	1.9 to 2.5
Second attendance	106 (43)	166 (13)	3.2	2.6 to 3.9
Third attendance	69 (28)	82 (6.6)	4.2	3.2 to 5.6
Signs				
Finger clubbing	11 (4.5)	1 (0.08)	55	7.1 to 420
Investigations				
Thrombocytosis	34 (14)	19 (1.5)	8.9	5.2 to 15
Abnormal spirometry	24 (9.7)	14 (1.1)	8.6	4.5 to 16

*Values shown as number (%).

studied. In seven cases a firm clinical diagnosis of cancer was made but initial biopsies were negative; positive histological results were obtained later. In these cases the date of diagnosis for study purposes was changed to the date of the first specialist investigation which was 36–119 days before histological proof.

Histological results were available for 237 of the 247 cases: 80 (32%) had squamous carcinoma, 57 (23%) adenocarcinoma, 52 (21%) small cell carcinoma, 21 (9%) large cell carcinoma, and 27 (11%) unspecified carcinoma. The remaining 10 cancers had been diagnosed clinically on strong radiological evidence. Staging data were available for only 134 (54%) of cases.

In obtaining the controls for study, 1417 were originally generated but 182 could not be used: 118 were ineligible (seven had previous lung cancer; 98 (7.4% of those available for study) had no consultations in the two year period; and 13 resided outside Exeter at diagnosis), and in 64 the notes were unobtainable (60 had left Exeter, four had died). For 221 cases, all age-matched controls were available within 1 year of the age of the case, for 17 cases within 2 years, and for the remaining nine cases within 4 years. These totals include 205 (83%) cases and 102 (8.3%) controls who had died at the time of study but whose notes were retrievable. Demographic details and the use of primary care by the subjects are shown in table 1.

For consultation and code measures over the whole 2 years, there was strong evidence of a higher rate of occurrence in cases than controls ($p<0.001$). Differences excluding the last 180 days were not significant ($p = 0.17$ for consultations, $p = 0.82$ for codes).

Quality of coding

Inter-observer variation in coding was examined by randomly selecting 188 codes. All four coders then coded the same records. The reliability coefficient was 0.83 (95% CI 0.75 to 0.90).[23]

Identification of independent associations with cancer

A total of 225 variables were recorded in 2.5% or more of either cases or controls. Selected univariable analyses are shown in table 2. All the variables in table 2 were significantly more common in cases than in controls ($p<0.001$). Smoking status was available for 1173 subjects (79.1%): 677 (45.7% of the total) were non-smokers, 204 (13.8%) were ex-smokers, and 292 (19.7%) current smokers. Platelet count was measured in 132 (53%) of the 247 cases and 34 (26% of these) were found to have thrombocytosis

Table 3 Multivariable analysis of the features of lung cancer

Variable	Odds ratio	95% CI	p value
Symptoms			
Loss of appetite*	86	3.6 to 2100	0.006
Haemoptysis	32	13 to 81	<0.001
Dyspnoea*	4.7	2.7 to 8.0	<0.001
Loss of weight	4.3	2.2 to 8.2	<0.001
Fatigue*	3.2	1.7 to 6.0	<0.001
Chest pain	2.9	1.8 to 4.7	<0.001
Second attendance with cough	2.7	1.7 to 4.4	<0.001
Sign			
Finger clubbing	18	1.7 to 190	0.016
Investigations			
Thrombocytosis	9.3	3.4 to 26	<0.001
Abnormal spirometry	7.5	2.8 to 21	<0.001
Smoking			<0.001
Non-smoker (reference category)	1.0		
Current smoker	9.7	5.3 to 18	
Ex-smoker	5.9	3.0 to 12	
Smoking status unknown	5.4	2.8 to 10	
Interaction terms			
Dyspnoea with fatigue	0.28	0.11 to 0.73	0.006
Loss of appetite in patients over 70 years	0.13	0.024 to 0.76	0.02

*Variables with interactions.

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.

Table 4 Multivariable results excluding the final 180 days

Variable	Odds ratio	95% CI	p value
Symptoms			
Haemoptysis	4.0	1.6 to 10	0.004
Dyspnoea	2.4	1.6 to 3.5	<0.001
Investigations			
Abnormal spirometry	3.5	1.2 to 9.6	0.02
Smoking			<0.001
Non-smoker	1.0		
Current smoker	9.0	5.8 to 14	
Ex-smoker	6.8	4.2 to 11	
Smoking status unknown	4.3	2.7 to 6.9	

(platelet count $>400\times10^9$/l). The median (interquartile range) time before diagnosis for these subjects was 60 (95% CI 36 to 203) days. In contrast, 396 of the 1235 controls (32%) had a platelet count measured with only 19 (5%) having thrombocytosis.

Multivariable analyses

From univariable conditional logistic regressions, 97 variables were considered for multivariable analyses. The first presentation with cough was not associated with lung cancer in the multivariable analyses, unlike the second presentation. The latter was therefore used in further multivariable modelling. Eleven variables remained in the final multivariable model (table 3). As an additional check, each of the discarded variables was added individually to the final model and none was associated with lung cancer. Sixteen clinically plausible interactions were tested in the final model. Two interactions were identified: (1) between dyspnoea and fatigue and (2) between loss of appetite and age below 70 years (table 3). There were no interactions with sex.

Timing of variable occurrence and analysis excluding the last 180 days

Multivariable analysis using data excluding the last 180 days was performed to identify early features of lung cancer (table 4).

The timings of presentations to primary care with haemoptysis, dyspnoea, and for abnormal spirometry results related to the date of diagnosis are shown in fig 1 which compares the monthly moving average number of presentations to primary care for each variable between cases and controls.

PPVs for a patient consulting a doctor in primary care

PPVs for lung cancer of selected variables individually, when paired with a second feature, and when the patient has presented with the same feature a second time are shown in fig 2. The variables chosen for fig 2 were those independently associated with lung cancer in the multivariable analysis, except for clubbing of the fingers where the numbers were too small for calculation of PPVs. Dyspnoea was rarely an isolated symptom: only 10 of the 139 cases with dyspnoea had no second symptom. PPVs were also calculated for two age strata (40–69 and 70+ years). All the variables in fig 2 except thrombocytosis had higher PPVs in older patients, reflecting the fourfold higher incidence of lung cancer in subjects aged over 70 years compared with those aged under 70. In older patients haemoptysis had a PPV of 7.1%, abnormal spirometry 4.2%, and the remaining variables in the range 0.9–2.2%. The only PPVs above 1% in patients aged 40–69 years were loss of appetite (1.1%) and thrombocytosis (3.0%). We also calculated PPVs for the subgroup of smokers and ex-smokers combined, and for non-smokers. In smokers and ex-smokers PPVs were approximately twice those for the study as a whole, and PPVs for non-smokers were one third to one half of those in the study as a whole.

DISCUSSION

Ten clinical features were found to be independently associated with the future development of lung cancer, as well as cigarette smoking. Three of these plus smoking

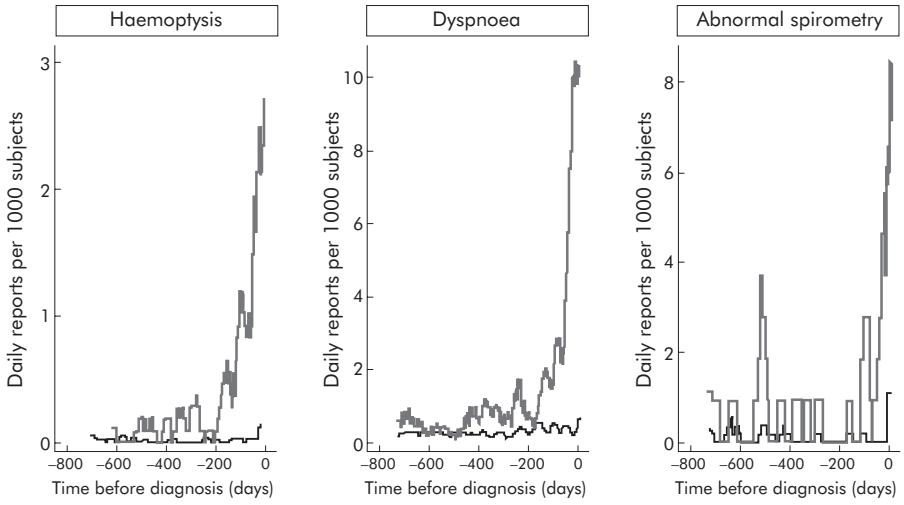


Figure 1 Timing of symptom presentation (haemoptysis, dyspnoea, and abnormal spirometry) to primary care in cases and controls. Time 0 is the date of diagnosis in cases. Grey = cases; black = controls. Note that the *y* axes have different scales.

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.

Cough	Fatigue	Dyspnoea	Chest pain	Loss of weight	Loss of appetite	Thrombo-cytosis	Abnormal spirometry	Haemoptysis	
0.40 0.3, 0.5	**0.43** 0.3, 0.6	**0.66** 0.5, 0.8	**0.82** 0.6, 1.1	**1.1** 0.8, 1.6	**0.87** 0.6, 1.3	**1.6** 0.8, 3.1	**1.6** 0.9, 2.9	**2.4** 1.4, 4.1	PPV as a single symptom
0.58 0.4, 0.8	**0.63** 0.5, 0.9	**0.79** 0.6, 1.0	**0.76** 0.6, 1.0	**1.8** 1.1, 2.9	**1.6** 0.9, 2.7	**2.0** 1.1, 3.5	**1.2** 0.6, 2.6	**2.0** 1.1, 3.5	Cough
	0.57 0.4, 0.9	**0.89** 0.6, 0.3	**0.84** 0.5, 1.3	**1.0** 0.6, 1.7	**1.2** 0.7, 2.1	**1.8**	**4.0**	**3.3**	Fatigue
		0.88	**1.2** 0.9, 1.8	**2.0** 1.2, 3.8	**2.0** 1.2, 3.8	**2.0**	**2.3**	**4.9**	Dyspnoea
			0.95 0.7, 1.4	**1.8** 1.0, 3.4	**1.8** 0.9, 3.9	**2.0**	**1.4**	**5.0**	Chest pain
				1.2 0.7, 2.3	**2.3** 1.2, 4.4	**6.1**	**1.5**	**9.2**	Loss of weight
					1.7	**0.9**	**2.7**	**>10**	Loss of appetite
							3.6	**>10**	Thrombocytosis
								>10	Abnomal spirometry
								17	Haemoptysis

Figure 2 Positive predictive values (PPVs) for lung cancer for individual risk markers and for pairs of risk markers in combination (against a background risk of 0.18%). Notes: (1) The bold figure in each cell is the PPV when both features are present and the two smaller figures represent the 95% confidence intervals for the PPV. These have not been calculated when any cell in the 2×2 table was below 10 (invariably this was because too few controls had both features). For three pairs of symptoms, no controls had the combination; while strictly speaking undefined, these PPVs must logically be very high and so they have been set as >10%. (2) The yellow shading is for pairs of symptoms with a PPV over 1%, the amber shading is when the PPV is above 2%, and the red shading is for PPVs above 5%. (3) The cells along the diagonal relate to the PPV when the same feature has been reported twice. Thus, the cough/cough intersect is the PPV for lung cancer when a patient has attended twice with cough. For a third presentation with cough the PPV was 0.77% (95% CI 0.54 to 1.1).

remained associated with cancer at least 180 days before diagnosis. Associations with lung cancer have been previously reported for these features in secondary care studies. However, the strength and independence of these associations with lung cancer has not previously been shown in unselected populations. Our results can guide doctors when to consider investigation in patients with a symptom or symptoms that could represent lung cancer.

Strengths and weaknesses of the study

This is the first study to examine all the prediagnostic features of lung cancer together. We were also able to study many more cases than in the one previous case series from primary care.[7] Furthermore, as every general practice in a well defined population participated, we could identify and study almost all eligible cases, allowing us to calculate PPVs.

The proportion of our cases confirmed by histology is high, and our methods will inevitably have led to a few cases being missed. Taken together, this suggests that those missed were more ill, as such patients are less likely to be subjected to invasive investigation. Nonetheless, the number missed will have been small and is unlikely to have influenced the results greatly. Using a dataset to select variables by multivariable analysis and then calculating univariable PPVs on the same data set carries a risk of overestimation of the PPVs. As every variable we found to be independently associated with cancer had previously been reported in the literature, we are confident that the list of variables selected was robust.

One weakness of the study is that recording of symptoms and signs may vary between general practices. This was less of an issue for test results as these were extracted directly from the laboratory report. Doctors may record symptoms more thoroughly if they consider lung cancer to be a possibility. If so, the PPVs in this study will have been overestimated. The matched design will have partly compensated for such variations in recording. However, matching can also be a weakness as the ability to study the matched variable directly is lost. The two major factors affecting primary care consultation rates are age and sex.[24] The final decision regarding matching was a careful balance between insufficient matching and overmatching.

Symptoms

All the symptoms shown to be associated with lung cancer had PPVs below 2%, except for haemoptysis. This reflects the high frequency of respiratory symptoms in the general population and illustrates the difficulty doctors have in selecting which patients require investigation. Haemoptysis

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.

was reported by 20% of cases and 1.5% of controls, giving a PPV of 2.4%. Such a relatively high PPV supports recommendations that all patients with haemoptysis should be offered a chest radiograph.[15] A second presentation with haemoptysis increased the PPV to 17%. The frequency of haemoptysis in this study is lower than reported in the one previous UK series of 22 cases in which nine (41%) had experienced the symptom.[8] All other figures come from secondary care, ranging from 21% to 35%.[11] There is no previous literature to allow comparison in controls, either from primary or secondary care.

Having a second symptom with haemoptysis increased the risk of cancer markedly. The single exception was cough, where haemoptysis with cough had a lower PPV (2.0%) than haemoptysis alone. A likely explanation for this is the common alternative diagnosis of respiratory infection in which cough may be accompanied by streaky haemoptysis.[25]

Haemoptysis was also associated with cancer after removal of the final 180 days. This could represent doctors failing to consider lung cancer as a possibility. Alternatively, investigations may have been negative or misleading. For example, chest radiographs can be negative in lung cancer.[12 26] In other chest radiographs an ill defined shadowing is reported which requires further investigation, although it is unlikely such a finding would lead to a delay in diagnosis of 180 days.

The remaining six symptoms (loss of appetite, loss of weight, dyspnoea, chest pain, fatigue and cough) individually posed a low risk for lung cancer. However, as with haemoptysis, when more than one symptom was present the risk of cancer rose. Of these six symptoms, only dyspnoea remained associated with lung cancer more than 180 days before diagnosis. This finding supports a retrospective study which reported that dyspnoea was the initial complaint in 17% of lung cancer patients[27] and an interview study in which patients reported symptoms of their cancer for a median of 12 months before diagnosis.[8] In the cases reported here, dyspnoea was rarely an isolated symptom. This accords with research from clinics for investigation of isolated dyspnoea which very rarely identified lung cancer.[28] This suggests that investigation of isolated dyspnoea should concentrate on non-malignant causes such as heart failure, and only if a second symptom is reported should lung cancer become the focus of investigation.

Cough is the most common symptom seen in primary care.[9] It is also the most common symptom in lung cancer, occurring in 65% of cases in this study. Re-attendance with cough was also very common in cases. The risk of lung cancer increased with each attendance, but still remained below 1%. Furthermore, no pair of symptoms with cough had a PPV over 2%. However, cough is the first symptom of cancer in nearly a quarter of patients, so it should not be readily dismissed as a predictor of lung cancer.[27] The low PPV for a first presentation of cough suggests that a doctor need not be particularly concerned about lung cancer in the absence of other suspicious symptoms. When a patient presents with cough, both doctor and patient can afford to wait a short time to allow the diagnosis to become clearer. It is highly unlikely that a delay of a few days in diagnosing lung cancer will have a material impact on the chance of survival. It is therefore reasonable to suggest a chest radiograph for a re-attendance with unexplained cough that has persisted for 3 weeks or more. This guidance would mean that some patients with a slow recovery from an upper respiratory infection would qualify for chest radiography, although in most of these patients the benign nature of the presentation would be clear.

Investigations

An abnormal spirometric test result was associated with lung cancer and, like dyspnoea, this finding remained after removal of the last 180 days. This accords with older studies which reported an association between chronic obstructive airways disease and cancer, independently of smoking.[29 30] It seems sensible to suggest that spirometric testing should be performed when there is no clear diagnosis in a dyspnoeic patient.

The association between thrombocytosis and lung cancer is an important finding. Although only 14% of cases had thrombocytosis, almost half had not been tested. The PPV for thrombocytosis was 1.6%. Furthermore, the platelet count was raised a median of 60 days before diagnosis. The only other study to link thrombocytosis with lung cancer examined patients admitted to hospital for investigation of suspected lung cancer; thrombocytosis was found in 53% of patients with lung cancer and in 8% of those without.[31] The platelet counts were taken during the admission to hospital, whereas in this study they had been taken in primary care some time before the diagnosis was made and, in some instances, before the cancer was even suspected. Serious consideration should be given to the possibility of lung cancer when thrombocytosis is found in a patient with respiratory symptoms.

Some symptoms and classical presentations of lung cancer were not identified in this study. These include hoarseness,[27] stridor, superior vena cava obstruction, or shoulder pain from a Pancoast tumour. Some of these presentations are rare and, despite the large number of cases, it is likely too few of these atypical presentations occurred. A second possible explanation is that hoarseness and abnormal investigation results such as hyponatraemia and a raised erythrocyte sedimentation rate—all of which were associated with cancer in the univariable analyses—are features of late cancer. In the multivariable analysis these variables were no longer independently associated with cancer in the presence of the other variables.

Until this study was performed, the decision about when to investigate a patient with possible lung cancer has had a very weak evidence base. With these results, clinicians have a guide to the risk of lung cancer when a patient presents with one or more symptoms. The PPVs give an initial guide when a single feature or pair of features is present. The implications of combinations of symptoms can be gleaned from the multivariable analysis. The results can be used by healthcare organisations to improve guidelines for selection of patients for rapid investigation, and also to inform the general public, with the caveat that they come from a study in patients who have already made the decision to consult their doctor. The one symptom patients do not delay reporting to their doctor is haemoptysis.[8] However, this is a relatively uncommon symptom of lung cancer and more benefit may be gained from educating the public about persistent cough and dyspnoea.

ACKNOWLEDGEMENTS

The authors thank all 21 general practices in Exeter, the Dendrite personnel, and the Patients and Practitioners Service Authority without which this project would not have been successful. They also thank three anonymous journal referees for their very helpful reviews.

Authors' affiliations

W Hamilton, T J Peters, D Sharp, Academic Unit of Primary Health Care, Department of Community Based Medicine, University of Bristol, Bristol BS8 1AU, UK
A Round, East Devon Primary Care Trust, Dean Clarke House, Southernhay, Exeter EX1 1PQ, UK

Project funding from the UK Department of Health. The funding source had no role in the study other than financial support. All authors had full access to all data and take final responsibility for publication. WH is

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.

funded through his research practice (Barnfield Hill, Exeter) and RCGP/BUPA and NHS fellowships.

Competing interests: none

The views expressed in the publication are those of the authors and not necessarily those of the Department of Health.

REFERENCES

1 **Cancer Research UK**. *CancerStats Monograph 2004*. London: Cancer Research UK, 2004.
2 **Jemal A**, Murray T, Ward E, *et al*. Cancer statistics, 2005. *CA Cancer J Clin* 2005;**55**:10–30.
3 **Mulshine JL**, Smith R. Lung cancer—2: Screening and early diagnosis of lung cancer. *Thorax* 2002;**57**:1071–8.
4 **Ford LG**, Minasian LM, McCaskill-Stevens W, *et al*. Prevention and early detection clinical trials: opportunities for primary care providers and their patients. *CA Cancer J Clin* 2003;**53**:82–101.
5 **Smith RA**, Cokkinides V, Eyre HJ. American Cancer Society guidelines for the early detection of cancer, 2005. *CA Cancer J Clin* 2005;**55**:31–44.
6 **Fergusson R**, Gregor A, Dodds R, *et al*. Management of lung cancer in South East Scotland. *Thorax* 1996;**51**:569–74.
7 **Mansson J**, Bengtsson C. Pulmonary cancer from the general practitioner's point of view. Experience from the health centre area of Kungsbacka, Sweden. *Scand J Primary Health Care Suppl* 1994;**12**:39–43.
8 **Corner J**, Hopkinson J, Fitzsimmons D, *et al*. Is late diagnosis of lung cancer inevitable? Interview study of patients' recollections of symptoms before diagnosis. *Thorax* 2005;**60**:314–9.
9 **Okkes IM**, Oskam SK, Lamberts H. The probability of specific diagnoses for patients presenting with common symptoms to Dutch family physicians. *J Fam Pract* 2002;**51**:31–6.
10 **Gallagher A**, Thomas J, Hamilton W, *et al*. The incidence of fatigue symptoms and diagnoses presenting in UK primary care from 1990 to 2001. *J R Soc Med* 2004;**97**:571–5.
11 **Hamilton W**, Sharp D. Diagnosis of lung cancer in primary care: a structured review. *Fam Pract* 2004;**21**:605–11.
12 **Colice GL**. Detecting lung cancer as a cause of hemoptysis in patients with a normal chest radiograph: bronchoscopy vs CT. *Chest* 1997;**111**:877–84.
13 **Summerton N**. Cancer recognition and primary care. *Br J Gen Pract* 2002;**52**:5–6.
14 **Metcalf JV**, Smith J, Jones R, *et al*. Incidence and causes of rectal bleeding in general practice as detected by colonoscopy. *Br J Gen Pract* 1996;**46**:161–4.
15 **Department of Health**. *Referral guidelines for suspected cancer*. London: Department of Health, 2000.
16 **Thiru K**, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ* 2003;**326**:1070–0.
17 **Hippisley-Cox J**, Pringle M, Cater R, *et al*. The electronic patient record in primary care: regression or progression? A cross sectional study. *BMJ* 2003;**326**:1439–43.
18 **WONCA**. *ICPC-2. International Classification of Primary Care*. Prepared by the Classification Committee of WONCA. Oxford: Oxford University Press, 1998.
19 **de Lusignan S**, Minmagh C, Kennedy J, *et al*. A survey to identify the clinical coding and classification systems currently in use across Europe. *Medinfo* 2001;**10**(Pt 1):86–9.
20 **Knottnerus JA**. *The evidence base of clinical diagnosis*. London: BMJ Books, 2002.
21 **Spiegelhalter D**, Thomas A, Best N, *et al*. *WinBUGS user manual*. Cambridge: MRC, 2003.
22 **StataCorp**. *Stata Statistical Software:Release 8.0*. College Station, TX: Stata Corporation, 2001.
23 **Streiner D**, Norman G. *Health measurement scales*. 3rd ed. Oxford: Oxford University Press, 2003.
24 **McCormick A**, Fleming D, Charlton J. *Morbidity statistics from general practice. Fourth national study 1991–1992*. London: HMSO, 1995.
25 **Kuyvenhoven MM**, Spreeuwenberg C, Touw-Otten FW. Diagnostic styles of general practitioners confronted with ambiguous symptoms. An exploratory study. *Scand J Primary Health Care* 1989;**7**:43–8.
26 **Kaminski J**. Frequency and causes of hemoptysis and role of bronchoscopy in patients with normal chest roentgenogram hospitalized in the Department of Physiopneumonology Silesian Medical University in the years 1961–1996. *Pneumonologia i Alergologia Polska* 2001;**69**:663–8.
27 **Koyi H**, Hillerdal G, Branden E. A prospective study of a total material of lung cancer from a county in Sweden 1997–1999: gender, symptoms, type, stage, and smoking habits. *Lung Cancer* 2002;**36**:9–14.
28 **Nielsen LS**, Svanegaard J, Wiggers P, *et al*. The yield of a diagnostic hospital dyspnoea clinic for the primary health care section. *J Intern Med* 2001;**250**:422–8.
29 **Tockman MS**, Anthonisen NR, Wright EC, *et al*. Airways obstruction and the risk for lung cancer. *Ann Intern Med* 1987;**106**:512–8.
30 **Skillrud DM**, Offord KP, Miller RD. Higher risk of lung cancer in chronic obstructive pulmonary disease. A prospective, matched, controlled study. *Ann Intern Med* 1986;**105**:503–7.
31 **Pedersen LM**, Milman N. Diagnostic significance of platelet count and other blood analyses in patients with lung cancer. *Oncol Rep* 2003;**10**:213–6.

Thorax: first published as 10.1136/thx.2005.045880 on 14 October 2005. Downloaded from http://thorax.bmj.com/ on March 29, 2024 by guest. Protected by copyright.