



Ades, A. E., Biswas, M., Welton, N. J., & Hamilton, W. (2014). Symptom lead time distribution in lung cancer: natural history and prospects for early diagnosis. *International Journal of Epidemiology*, 43(6), 1865-1873. [dyu174]. <https://doi.org/10.1093/ije/dyu174>

Publisher's PDF, also known as Version of record

Link to published version (if available):  
[10.1093/ije/dyu174](https://doi.org/10.1093/ije/dyu174)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



---

Original article

## Symptom lead time distribution in lung cancer: natural history and prospects for early diagnosis

Anthony E Ades,<sup>1\*</sup> Mousumi Biswas,<sup>1</sup> Nicky J Welton<sup>1</sup> and William Hamilton<sup>2</sup>

<sup>1</sup>University of Bristol, School of Social and Community Medicine, Bristol, UK and <sup>2</sup>University of Exeter Medical School, Exeter, UK

\*Corresponding author. University of Bristol, School of Social and Community Medicine, 39 Whatley Road, Bristol BS8 2PS, UK. E-mail: t.ades@bristol.ac.uk

Accepted 31 July 2014

### Abstract

**Background:** Before their diagnosis, patients with cancer present in primary care more frequently than do matched controls. This has raised hopes that earlier investigation in primary care could lead to earlier stage at diagnosis.

**Methods:** We re-analysed primary care symptom data collected from 247 lung cancer cases and 1235 matched controls in Devon, UK. We identified the most sensitive and specific definition of symptoms, and estimated its incidence in cases and controls prior to diagnosis. We estimated the symptom lead time (SLT) distribution (the time between symptoms attributable to cancer and diagnosis), taking account of the investigations already carried out in primary care. The impact of route of diagnosis on stage at diagnosis was also examined.

**Results:** Symptom incidence in cases was higher than in controls 2 years before diagnosis, accelerating markedly in the last 6 months. The median SLT was under 3 months, with mean 5.3 months [95% credible interval (CrI) 4.5–6.1] and did not differ by stage at diagnosis. An earlier stage at diagnosis was observed in patients identified through chest X-ray originated in primary care.

**Conclusions:** Most symptoms preceded clinical diagnosis by only a few months. Symptom-based investigation would lengthen lead times and result in earlier stage at diagnosis in a small proportion of cases, but would be far less effective than standard screening targeted at smokers.

**Key words:** Lung cancer, lead time, symptom lead time, stage at diagnosis

---

### Key Messages

- Individuals with undiagnosed lung cancer attend general practice with suggestive but non-specific symptoms more often than matched controls, but the great majority of the excess attendances precede the diagnosis date by a only few months.
- As a result, a strategy of diagnostic imaging of patients with such symptoms will bring forward diagnosis, and will lead to earlier stage diagnosis, but the effects will be modest.
- A standard screening programme for smokers would be likely to have much greater benefits for those diagnosed, in terms of lead time and the proportion diagnosed at an early stage.

## Introduction

There has been increasing emphasis on the need for earlier investigation of suspected cancer in primary care. In the 2 years before diagnosis, individuals with lung cancer present more frequently in primary care than do matched controls with cancer-related symptoms.<sup>1–3</sup> Similar findings have been reported for colorectal,<sup>4</sup> ovarian<sup>5,6</sup> and other cancers.<sup>4</sup> These studies have given rise to optimism that earlier diagnosis could be achieved by instituting explicit criteria for investigation in primary care.<sup>7</sup> In the UK, the Department of Health launched early awareness campaigns in 2011, advising individuals with particular symptoms to attend their general practitioner (GP) surgery.<sup>8</sup> In several countries, including England, Wales, Scotland and Denmark, guidance recommends urgent cancer investigation in certain scenarios.

However, information on the predictive value and relative frequency of symptoms in cases and controls<sup>1–5</sup> does not allow one to quantify how much earlier diagnosis could be made if criteria for investigation were changed, nor how many cancers currently diagnosed at a late stage could be diagnosed at an earlier stage, when better treatment options exist. This study shows how this information can be derived and presents results for lung cancer. The same methods can be applied to other cancers, to provide insights into natural history before diagnosis, and to assess the role of symptom-based investigation.

The target parameter is the symptom lead time distribution (SLT), defined here as the interval between the occurrence of a particular symptom, or set of symptoms, attributable to cancer and the eventual diagnosis of cancer if no further action is taken. Some diagnoses of lung cancer are already the direct result of symptom presentation in primary care, because GPs already investigate by chest X-ray (CXR) on the basis of symptoms. In these patients, if no CXR had been requested, the date of diagnosis would have been later, and the SLT longer. This needs to be considered when estimating the SLT.

An investigational strategy based on symptoms presenting in primary care may lead to earlier diagnosis and better outcomes. However, the sensitivity and specificity of these

symptoms may approximate the sensitivity and specificity of smoking as a marker of lung cancer risk. We compare symptom-based and smoking-based investigational strategies in the light of our results.

## Methods

This paper presents a new analysis of the Cancer Prediction in Exeter (CAPER) study.<sup>1</sup> Data was collected on 247 lung cancer cases diagnosed in Exeter, England, 1998–2002, and 1235 age-, sex- and practice-matched controls. GP attendances and symptoms were recorded over the 2 years before the case was diagnosed. The cancer stage notation (TNM) at diagnosis was available from the local cancer registry, supplemented by searches of the primary care records where necessary. CXRs ordered by the GP, classified as ‘normal’, ‘abnormal’ and ‘suspected lung cancer’ were also recorded,<sup>9</sup> as were emergency admissions that led to a diagnosis of lung cancer.

## Symptoms

We examined attendances for cough, dyspnoea, chest pain, fatigue, loss of weight, loss of appetite and haemoptysis. From these we derived several sets of potential criteria for investigation by appropriate imaging. These were: (i) cough; (ii) dyspnoea; (iii) chest pain; (iv) fatigue, loss of weight, loss of appetite or haemoptysis; and (v) any of these seven symptoms. We also explored (vi) any two symptoms, including repeat attendances with the same symptom, presenting within 1, 2, 3, 4, 5 or 6 months, and (vii) the same six variants but excluding repeats of the same symptom. We used receiver operating characteristic (ROC) plots,<sup>10</sup> to assess the ability of symptom criteria to discriminate between cases and controls.

## Symptom incidence and symptom lead time distribution

We estimated the rate of symptom presentation in cases and controls during each quarter year before diagnosis

via Poisson regression, and interpreted the difference between cases and controls in quarter  $T$  as the rate attributable to lung cancer. The SLT distribution is given by the ratio of the cancer-related incidence of symptoms in quarter  $T$  to the sum of cancer-related incidence across all quarters. We also examined the proportion of all symptoms that were cancer-related, estimated as the incidence of symptoms in cases minus incidence in matched controls, divided by the incidence in cases. A Bayesian Markov chain Monte Carlo analysis was adopted, as it allows flexibility to estimate non-standard models. Full details of the statistical model with an example dataset and annotated programming code are provided as appendices available in [Supplementary data at IJE online](#).

SLT was estimated in (i) the full dataset and (ii) excluding cases who had a 'suspected lung cancer' finding on CXR in the 4 months before diagnosis, and who were not admitted as an emergency. These provided minimum and maximum estimates of SLT, respectively, as explained below.

### Ethics approval

Ethical approval for the original study was given by the North and East Devon research ethics committee.

## Results

### Stage at diagnosis

Of the 247 cases, 113 (45.7%) were reported as unknown stage at diagnosis. These were combined with the 55 diagnosed as Stage 4. On this basis, the proportions in Stages 1, 2, 3 and 4 are therefore 12.6%, 6.9%, 12.6% and 68%, respectively.

### Ability of alternative criteria to discriminate cases and controls

[Table 1](#) shows the proportion of cases and controls who ever presented with each symptom combination over the 2-year period. As previously reported,<sup>1–3</sup> symptoms vary in their sensitivity and specificity for lung cancer: isolated cough picks up over half the cases, but would also lead to investigations in 29.5% of controls. Haemoptysis is highly specific (present in only 1.5% of controls), but would identify only 20% of cases. At the other extreme, 'any symptom' is very sensitive (91.9%) but also very non-specific (47%) ([Figure 1](#)).

'Any two symptoms within 3 months' was the most discriminating criterion, based on the Youden Index (sensitivity+specificity-1) ([Table 1](#)), with 80.6% sensitivity and 23.2% false positives. All analyses reported below are based

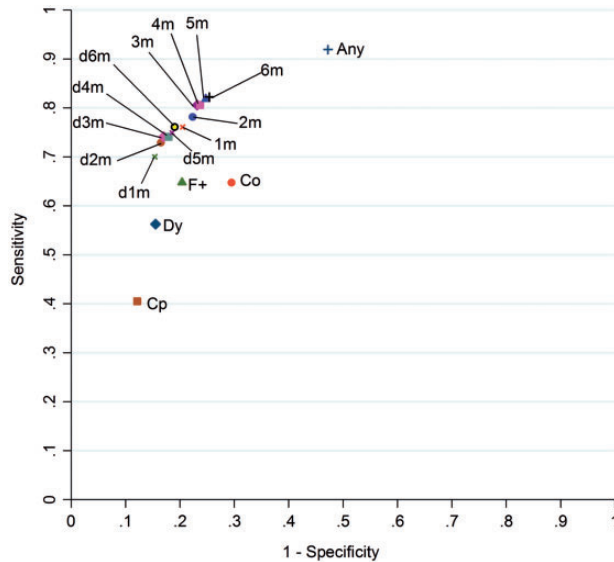
**Table 1.** Percentage of cases and controls who ever presented with symptoms meeting investigative criteria during the 2 years (Stage 4 includes cases with missing stage)

Symptom	Stage I N = 31	Stage II N = 17	Stage III N = 31	Stage IV N = 168	All cases N = 247	Controls N = 1235	Youden Index,% (95% CrI)
Cough	54.8	64.7	61.3	67.3	64.8	29.5	35.3 (29–49)
Dyspnoea	38.7	41.2	64.5	59.5	56.3	15.5	40.7 (34–47)
Chest pain	35.5	47.1	38.7	41.1	40.5	12.1	28.3 (22–35)
Fatigue	41.9	11.8	35.5	36.3	35.2	15.1	20.2 (14–26)
Loss of weight	12.9	41.2	22.6	29.2	27.1	4.4	22.8 (17–28)
Loss of appetite	12.9	17.6	25.8	19.0	19.0	4.0	15.1 (10–20)
Haemoptysis	29.0	29.4	12.9	19.0	20.2	1.5	18.7 (14–24)
Fatigue, loss of weight or appetite, haemoptysis	67.7	58.8	61.3	65.5	64.8	20.4	44.4 (38–51)
Any symptom	90.3	88.2	93.5	92.3	91.9	47.2	44.7 (40–49)
Two symptoms in 1 month	67.7	76.5	80.6	76.8	76.1	20.5	55.6 (50–61)
Two symptoms in 2 months	71.0	82.4	80.6	78.6	78.1	22.3	55.8 (50–61)
Two symptoms in 3 months	74.2	82.4	83.9	81.0	80.6	23.2	57.4 (52–63)
Two symptoms in 4 months	74.2	82.4	83.9	81.0	80.6	23.7	56.8 (51–62)
Two symptoms in 5 months	77.4	82.4	83.9	82.1	81.8	24.8	57.0 (52–62)
Two symptoms in 6 months	80.6	82.4	83.9	82.1	82.2	25.4	56.8 (51–62)
Two different symptoms in 1months	61.3	70.6	74.2	70.8	70.0	15.4	54.7 (49–61)
Two different symptoms in 2months	64.5	76.5	77.4	73.2	72.9	16.5	56.4 (50–62)
Two different symptoms in 3months	67.7	76.5	80.6	73.8	74.1	17.0	57.1 (51–63)
Two different symptoms in 4months	67.7	76.5	80.6	73.8	74.1	17.9	56.2 (50–62)
Two different symptoms in 5months	71.0	76.5	80.6	74.4	74.9	18.5	56.4 (51–62)
Two different symptoms in 6months	74.2	76.5	80.6	75.6	76.1	19.0	57.1 (51–63)

on this symptom criterion. Criteria based on two different symptoms lie on a similar ROC curve, with both the sensitivity and the false-positive rate reduced by 4–6% compared with criteria including repeat attendances with the same symptom. The prevalence of symptoms before diagnosis was not strongly correlated with stage at diagnosis.

### GP referrals for CXR

In total 167 (67.7%) of cases and 101 (8.2%) of controls were referred for CXR by their GP. Nearly every finding of



**Figure 1.** Sensitivity and specificity of symptom definitions in the CAPER study. Co, cough; m, months; Dy, dyspnoea; Cp, chest pain; F+, fatigue, loss of weight, loss of appetite, or haemoptysis; Any, any of the seven symptoms; nm, any two symptoms in *n* months; dnm, any two different symptoms in *n* months.

‘suspected lung cancer’ was associated with a diagnosis within 4 months (Table 2). However, only 51 of 89 (57%) ‘abnormal, but not suspicious of cancer’ CXR findings occurred within 4 months of diagnosis. In Table 3 we have grouped together as ‘diagnosis not influenced by CXR’ all those who either had no ‘suspected lung cancer’ or ‘abnormal’ CXR finding in the 4 months before diagnosis or who, regardless of CXR findings, had an emergency admission. These are, in effect, the cases whose date of diagnosis has not been materially brought forward by a CXR. Those with ‘suspected lung cancer CXR’ in the last 4 months had a symptom rate over the first 20 months of 0.21 per year, similar to the symptom rate in controls of 0.18. In contrast, those with an ‘abnormal but not suspicious of cancer’ CXR in the last 4 months have a symptom rate in months 1–20 (0.51) that is indistinguishable from the symptom rate in the ‘diagnosis not influenced by CXR’ cases (0.54). The two groups were therefore combined.

### Estimates of incidence and symptom lead time distribution

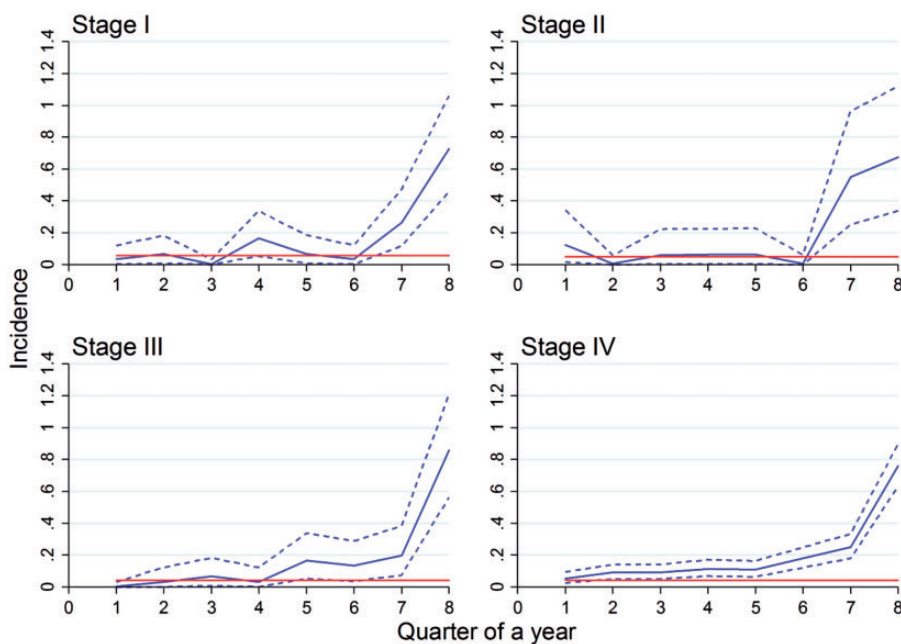
For all stages, the incidence of symptoms in cases rises from a level slightly higher than in controls, and accelerates rapidly in the 6 months before diagnosis (Figure 2). SLT was estimated in two datasets, first in the full dataset, and second in a dataset excluding those with a ‘suspected lung cancer’ CXR finding in the 4 months before diagnosis and no emergency admission. Figure 3 shows the three sets of SLTs for each stage: the minimum SLT, based on the entire dataset; the maximum SLT, based on the second dataset; and our preferred estimate of SLT, which is halfway between. The differences between the minimum and

**Table 2.** Numbers (rates x 100) of CXRs, and person-years at risk (PYARs), in cases and controls: timing and results

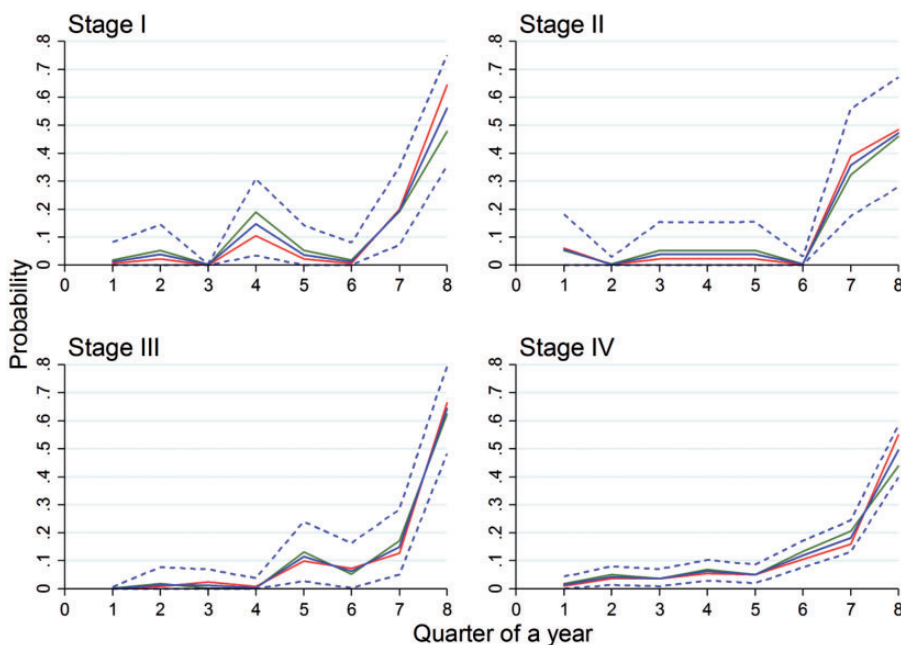
Patients	Interval, months	N	PYAR	‘Abnormal’ CXR	‘Suspected lung cancer’ CXR
Cases	1–20	247	411.7	38 (9.2)	2 (0.49)
Cases	21–24	247	8.23	51 (620)	101 (1227)
Controls	1–24	1235	2470	41 (1.66)	3 (0.12)

**Table 3.** Numbers of times the symptom criterion ‘any 2 symptoms in 3 months’ was met, and the rates; estimates as events divided by person-years at risk (PYARs), in cases and controls, by CXR findings in the last 4 months before diagnosis

Patients	Interval, months	Number of patients	PYAR	Number of events (rates, per year)
Cases, CXR ‘abnormal’	1–20	29	48.3	25 (0.51)
	21–24	29	7.25	37 (3.8)
Cases, CXR ‘suspected lung cancer’	1–20	92	153	33 (0.21)
	21–24	92	3.1	85 (2.75)
Cases ‘diagnosis not influenced by CXR’	1–20	126	210	115 (0.54)
	21–24	126	42	91 (2.15)
Controls	1–24	1235	2470	438 (0.18)



**Figure 2.** Incidence of ‘any two symptoms in 3 months’ in cases and controls. Blue lines represent cases, dashed blue lines are 95% credible intervals, red lines are controls.



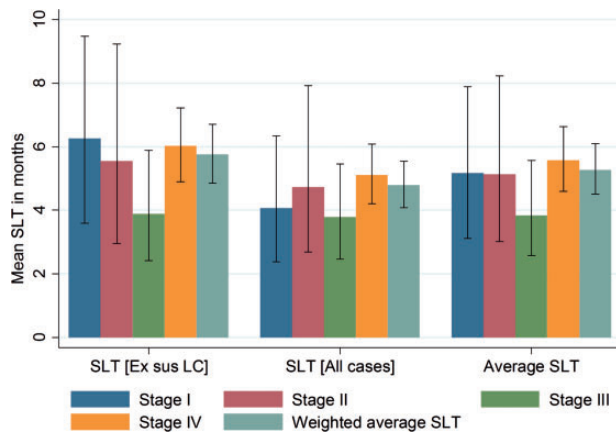
**Figure 3.** Symptom lead time distribution, ‘any two symptoms in 3 months’. Red line represents full dataset, green line excludes ‘suspected lung cancer’ CXR findings in the last 4 months, blue line intermediate or average, dashed blue lines 95% credible intervals.

maximum results are relatively narrow. Mean SLTs by stage at diagnosis are shown in Figure 4. The average SLTs, weighted in proportion to the sample stage distribution, are 4.8 months (minimum estimate), 5.7 months (maximum) and 5.3 months [95% CrI: 4.5–6.1] as the intermediate estimate. These figures represent the SLT of symptoms attributable to lung cancer in the 81 % of cases who met the

criterion ‘any two symptoms within the same 3 months’. Symptoms attributable to lung cancer comprised 79% [95% CrI: 76%–82%] of all symptoms in the 2 years before diagnosis. Overall, 52% of cases had SLTs of 3 months or less, suggesting a median SLT of just under 3 months.

Very similar SLT distributions were found for other investigative criteria, such as cough, any symptom and ‘any





**Figure 4.** Mean symptom lead time (SLT), 'any two symptoms in 3 months' by stage, including and excluding (Ex sus LC) cases with 'suspected lung cancer' findings on CXR in the last 4 months. Average SLT is intermediate.

two different symptoms in 3 months'. Incidence and SLT plots are shown in the appendices available in [Supplementary data at IJE online](#).

## Discussion

This paper joins a series of reports of higher risks of early symptom presentation in cancer cases compared with matched controls,<sup>1–5</sup> which have given rise to optimism that cancers could be diagnosed earlier, and at an earlier stage, if symptom-based criteria for immediate investigation were adopted in primary care. In the UK, the Department of Health has advocated Early Awareness campaigns (see: <http://www.nhs.uk/be-clear-on-cancer/Pages/beclearoncancer.aspx>), encouraging patients with symptoms, such as persistent cough for lung cancer, or haematuria for bladder cancer, to attend general practice.<sup>8</sup> The lung cancer campaign has been reported as increasing GP attendance rates for cough symptoms, the number of cancers diagnosed and the proportion diagnosed at an early stage.<sup>11</sup>

Symptom rates and CXR rates in the study on which the present analysis is based accord with much earlier larger studies.<sup>2,3</sup> An ideal symptom-based strategy would be based on symptom criteria that are both sensitive and specific markers of lung cancer. Our results suggest that isolated symptoms are inferior to pairs of symptoms or repeat symptoms. The most discriminating criterion was any two symptoms presenting within 3 months. In the 2-year study window, this identified 81% of cases and 23% of controls. In the UK, the Be Clear on Cancer Campaign uses a criterion of persistent cough,<sup>11</sup> which is less sensitive and less specific. Current National Institute for Health and Care Excellence (NICE) guidance<sup>7</sup> lists only single symptoms, without reference to paired or multiple symptoms,

although the clinician is expected to exercise judgment. Similar guidelines exist in Scotland. In Denmark, criteria for investigation are entirely at GPs' discretion.

Much larger studies have shown that sensitivity and specificity of investigative criteria can be improved by adding smoking, age, chronic obstructive pulmonary disease, pneumonia and family history of cancer,<sup>2,3</sup> and a number of algorithms have been suggested including the Liverpool Lung Cancer Risk Model.<sup>12</sup> However, sensitivity analyses carried out on the CAPER data, but not shown here, demonstrate that neither smoking nor age impact on SLT.

The SLT distribution<sup>6</sup> is the critical information required to inform policies on investigational strategies. We defined SLT as the time between symptoms attributable to cancer and clinical diagnosis if no action is taken. One contribution of this paper is that it shows how to estimate SLT from case-control data, taking into account the fact that GPs already order CXR following symptoms. Exclusion of symptoms occurring in the last 3 months before diagnosis<sup>1,3,6</sup> gives a more realistic impression of the sensitivity of symptom criteria that would be obtained in practice<sup>6</sup> (see below), but removes a large proportion of symptoms from the denominator, biasing lead times upwards. Further, in patients diagnosed because of symptoms, lead times associated with all symptoms, not just those in the last 3 months, are censored and therefore underestimated. Our approach was to estimate SLT both including and excluding cases whose CXR finding of 'suspected lung cancer' could have advanced the date of diagnosis, generating lower and upper estimates of SLT, respectively, and taking the intermediate as our preferred estimate. On this basis we estimated the mean SLT to be 5.3 months (95% CrI: 4.5–6.1) for lung cancer.

Information on the pre-diagnosis natural history of cancers centres on the mean sojourn time, that is the time spent with detectable cancer pre-diagnosis. Negatively correlated estimates of mean sojourn time and test sensitivity can be jointly derived from screening trials with two or more screening rounds.<sup>13</sup> Estimates of mean sojourn time in lung cancer have ranged from 1.38 to 3.92 years.<sup>14–16</sup> The latter estimate is to be preferred because it was accompanied by an estimated CXR sensitivity of 0.57, which accords closely with estimates from the National Lung Screening Trial.<sup>17</sup> More precise estimates of sojourn time from this landmark trial are yet to be published.

The results presented here confirm that many sojourn times are well in excess of 2 years, and provide a new, complementary, perspective on the natural history of lung cancer before diagnosis. A strength of the study was the inclusion of information on staging: although this was incomplete, the stage distribution in the study was in good agreement with the 9.4%, 4.7%, 20% and 61.9% in stages

1,2,3, and 4, respectively, and missing, reported from the two UK registries with the most complete stage data 2004–07.<sup>18</sup>

The main finding was that the majority of symptom presentations in primary care occur in the last 6 months before diagnosis. SLT was largely independent of TNM stage at diagnosis. This is a crucial finding as it suggests that patients diagnosed at stage 4 were not presenting with cancer-related symptoms at a higher rate while they were in stage 1, otherwise longer SLTs would be associated with later stage diagnosis. Instead, it seems that, regardless of stage at diagnosis, the occurrence of most symptoms is late and reflects an accelerating process during the final 3–6 months, leading eventually to diagnosis, whether as a result of GP-ordered investigation, accidental discovery by other specialists or later presentation as an emergency.

Overlaid on this process, however, is a robust finding that a small proportion of patients present with symptoms associated with cancer much earlier. Diagnosis could therefore be advanced considerably, with a probable improved staging at diagnosis, but only in a minority of cases: only 29% of all symptoms attributable to lung cancer occurred more than 6 months before diagnosis, and 19% of cases did not present these symptoms at all in the 2 years before diagnosis.

The most serious limitation of the study is that observations only run to the 2 years before diagnosis, and may therefore underestimate the SLT. However, if the rate of symptom presentation 25–36 months before diagnosis is at the same rate as that we observed in the 7–24 months, the average lead time would rise from 5.3 to 8.3 months. Note that these SLT estimates overestimate the lead time to start of treatment. Another limitation is the small study size, which prevents us extracting the more detailed information on time between symptom presentations that would be required to predict the performance of a sustained programme of symptom-based investigation.

## Implications for investigative strategies

We can compare the benefits of a programme of investigation based on symptoms presenting in primary care, with a standard screening programme targeted at smokers. The mean sojourn time, about 4 years for lung cancer, represents an upper bound on the lead time that can be obtained from a standard screening programme. For common sojourn time distributions, in a prevalent screen with a 100% sensitive test, the expected lead time is at least one-half of the mean sojourn time, in other words over 2 years. Recent UK data suggest that 66% of lung cancers can be found in the 31.3% most smoking-exposed population,<sup>3</sup> giving an expected lead time of at least 2 years for 66% of cancers and zero for the remaining 34, an average of over 1.32 years.

We found that 81% of lung cancers met the symptom criteria at least once, and 23% of controls did so. However, the case-control design aligns the cancer cases so they all present at their highest rate during the study. In a symptom-based programme run over a random 2-year period, in which patients were tested no more than once, a far lower sensitivity would be obtained, and lower specificity would be obtained if the programme ran for more than 2 years.

Some 21% of the cases would be picked up serendipitously in the sense that their symptoms were not cancer-related. The patients concerned then benefit from a substantial lead time. However, this detracts from efficiency and yield as it is equivalent to screening a proportion of patients at random. Taking this into account, and optimistically assuming symptoms would identify 81% of cases, the mean lead time would be 0.62 years, or 0.78 years if cases present symptoms at a higher rate over a 36-month period before diagnosis. During a sustained programme, whether symptom- or smoking-based, lead times will be higher as cases will be detected earlier, but whatever interval is set between consecutive screens, lead times

**Table 4.** Stage distribution at diagnosis. Patients with and without a CXR finding of ‘suspected lung cancer’ in the CAPER study, compared with routine data Eastern Region, UK, National Cancer Intelligence Network (NCIN), 2006–08; and with the National Lung Screening Trial (NLST) 1st (prevalent) round<sup>19</sup> and 2nd and 3rd (incident) screening rounds<sup>17</sup>

Study		% Stage 1	% Stage 2	% Stage 3	% Stage 4, or missing	Total cases (100%)
CAPER	‘Suspected LC’ on CXR	15.5	7.8	13.6	63.1	103
	No ‘suspected LC’ on CXR	9.7	6.3	11.8	71.5	144
NCIN	GP route of referral	14.7	6.5	29.4	59.4	3977
	Other route of referral	9.3	3.5	23.2	63.9	5695
NLST	Prevalent, CXR	46.5	8.3	25.6	22.7	136
	Prevalent, CT	58.8	16.7	29.6	16.9	270
	Incident, CXR	48.3	9.8	19.6	22.4	143
	Incident, CT	64.6	7.4	14.8	13.2	379

LC, lung cancer; CT, computerized tomography.



will be shorter in the symptom-based programme, as cases will have met symptom criteria before they are investigated, which occurs relatively late.

These pessimistic conclusions about lead times in symptom-based vs smoking-based strategies are supported by data on the distribution of stage at diagnosis. Table 4 shows a useful, but modest increase in the proportion of stage 1 diagnoses in patients receiving a 'suspected lung cancer' finding on CXR in the CAPER study, very similar to unpublished data on route of diagnosis from the National Cancer Intelligence Network. However, the extent of the shift towards stage 1 due to symptom-based diagnosis in primary care is dwarfed (Table 4) by what has been achieved in the National Lung Screening Trial.<sup>17,19</sup> However, a full evaluation would need to take account of the over-diagnosis in lung cancer screening.<sup>20,21</sup>

Lung cancer is unusual in that screening can be targeted on a relatively sensitive and specific marker, smoking. In other cancers a symptom-based strategy based on sensitive and specific criteria could be effective, and cost-effective, compared with no intervention, or as an adjunct to screening. However, a recent study of ovarian cancer, to our knowledge the only other study discussing SLT, concluded that a symptom-based programme could advance diagnosis by no more than 3 months.<sup>6</sup>

Patients presenting in primary care require diagnosis and treatment, and guidance on criteria for investigation is required. However, whether increased symptom-based investigation can lead to better cancer outcomes can only be assessed by study of the SLT distribution, based on analyses of large-scale case-control studies, looking at primary care data over a long period before diagnosis.

## Supplementary Data

Supplementary data are available at *IJE* online.

## Funding

This work was supported by the National Institute for Health Research (NIHR) Programme Grants for Applied Research Programme, RP-PG-0608–10045. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funders of this study had no role in: study design; data collection, analysis, and interpretation; or the writing of this research paper. The corresponding author had full access to all data and final responsibility for the decision to submit for publication.

## Acknowledgments

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); Clare Bankhead; Alison Clutterbuck; Jon Emery; Ardiana Gjini; Joanne

Hartland; Maire Justice; Jenny Knowles; Helen Morris; Richard Neal; Peter Rose; and Greg Rubin.

**Conflicts of interest:** W.H. is the clinical lead for the ongoing revision of the NICE 2005 guidance. His contribution to this article is in a personal capacity, and is not to be interpreted as representing the view of the Guideline Development Group, or of NICE itself. The remaining authors have no conflict of interest relevant to this research paper.

## References

1. Hamilton W, Peters T, Round A, Sharp D. What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. *Thorax* 2005;**60**:1059–65.
2. Hippisley-Cox J, Coupland C. Identifying patients with suspected lung cancer in primary care: Derivation and validation of an algorithm. *Br J Gen Pract* 2011;**61**:715–23.
3. Iyen-Omofoman B, Tata L, Baldwin D, Smith C, Hubbard R. Using socio-demographic and early clinical features in general practice to identify people with lung cancer earlier. *Thorax* 2013;**68**:451–59.
4. Hamilton W. The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. *Br J Cancer* 2009;**101**(Suppl 2):S80–86.
5. Friedman GD, Skilling JS, Udaltsova NV, Smith LH. Early symptoms of ovarian cancer: a case-control study without recall bias. *Fam Pract* 2005;**22**:548–53.
6. Lim AWW, Mesher D, Gentry-Maharaj A *et al*. Predictive value of symptoms for ovarian cancer: comparison of symptoms reported by questionnaire, interview, and general practitioner notes. *J Natl Cancer Inst* 2012;**104**:1–11.
7. National Institute for Health and Care Excellence (NICE). *Referral Guidelines for Suspected Cancer*. London: NICE, 2005.
8. Department of Health. *Promoting Early Diagnosis of Breast, Bowel and Lung Cancers*. London: DH, 2012.
9. Stapley S, Sharp D, Hamilton W. Negative chest x-rays in primary care patients with lung cancer. *Br J Gen Pract* 2006;**56**: 570–73.
10. Green DM, Swets JA. *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.
11. Cancer Research UK. *Be Clear on Cancer: Evaluation Summary May 2014*. London: Cancer Research UK, 2014.
12. Cassidy A, Myles J, van Tongeren M *et al*. The LLP risk model: an individual risk prediction model for lung cancer. *Br J Cancer* 2008;**98**:270–76.
13. Prevost TC, Launoy G, Duffy SW, Chen HH. Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. *Am J Epidemiol* 1998;**148**: 609–19.
14. Jang H, Kim S, Wu D. Bayesian lead time estimation for the Johns Hopkins lung project data. *J Epidemiol Global Health* 2013;**3**:157–63.
15. Chien C, Chen T. Mean sojourn time and effectiveness of mortality reduction for lung cancer screening with computed tomography. *Int J Cancer* 2008;**122**:2594–99.
16. Wu D, Erwin D, Rosner GL. Sojourn time and lead time projection in lung cancer screening. *Lung Cancer* 2011;**72**:322–26.

17. Aberle DR, DeMello S, Berg CD *et al.* Results of the two incidence screenings in the national lung screening trial. *N Engl J Med* 2013;**369**:920–31.
18. Walters S, Maringe C, Coleman M *et al.* Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: A population-based study, 2004–2007. *Thorax* 2013;**68**:551–64.
19. Church TR, Black WC, Aberle DR *et al*; National Lung Screening Trial Research Team. Results of initial low-dose computed tomographic screening for lung cancer. *N Engl J Med* 2013;**368**:1980–91.
20. Marcus P, Bergstralh E, Zweig M, Harris A, Offord K, Fontana R. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. *J Natl Cancer Inst* 2006;**98**:748–56.
21. Patz EJ, Pinsky P, Gatsonis C *et al.* Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA* 2014; **174**:269–74.