



Biswas, M., Ades, A. E., & Hamilton, W. T. (2015). Symptom lead times in lung and colorectal cancers: what are the benefits of symptom-based approaches to early diagnosis? *British Journal of Cancer*, 112, 271-277. <https://doi.org/10.1038/bjc.2014.597>

Peer reviewed version

Link to published version (if available):
[10.1038/bjc.2014.597](https://doi.org/10.1038/bjc.2014.597)

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

© 2015 Cancer Research UK

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/>

Keywords: Symptom lead time; Symptom-based diagnosis; lung cancer; colorectal cancer

Symptom lead times in lung and colorectal cancers: what are the benefits of symptom-based approaches to early diagnosis?

M Biswas¹, A E Ades^{*,1} and W Hamilton²

¹*School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK and*

²*Institute of Health Service Research, University of Exeter Medical School, College House, St Luke's Campus, Exeter EX2 4TE, UK*

Background: Individuals with undiagnosed lung and colorectal cancers present with non-specific symptoms in primary care more often than matched controls. Increased access to diagnostic services for patients with symptoms generates more early-stage diagnoses, but the mechanisms for this are only partially understood.

Methods: We re-analysed a UK-based case-control study to estimate the Symptom Lead Time (SLT) distribution for a range of potential symptom criteria for investigation. Symptom Lead Time is the time between symptoms caused by cancer and eventual diagnosis, and is analogous to Lead Time in a screening programme. We also estimated the proportion of symptoms in lung and colorectal cancer cases that are actually caused by the cancer.

Results: Mean Symptom Lead Times were between 4.1 and 6.0 months, with medians between 2.0 and 3.2 months. Symptom Lead Time did not depend on stage at diagnosis, nor which criteria for investigation are adopted. Depending on the criteria, an estimated 27–48% of symptoms in individuals with as yet undiagnosed lung cancer, and 12–32% with undiagnosed colorectal cancer are *not* caused by the cancer.

Conclusions: In most cancer cases detected by a symptom-based programme, the symptoms are caused by cancer. These cases have a short lead time and benefit relatively little. However, in a significant minority of cases cancer detection is serendipitous. This group experiences the benefits of a standard screening programme, a substantial mean lead time and a higher probability of early-stage diagnosis.

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 cancer present more frequently in primary care than matched controls with non-specific but suggestive symptoms. This has been reported for lung (Hamilton *et al*, 2005; Hippisley-Cox and Coupland, 2011; Iyen-Omofoman *et al*, 2013), colorectal (Hamilton, 2009), and ovarian cancers (Friedman *et al*, 2005; Hamilton, 2009; Lim *et al*, 2012). Similar findings have been published 1 year before diagnosis in pancreatic, oesophago-gastric, bladder, renal, and uterine cancers (Shephard *et al*, 2012, 2013; Stapley *et al*, 2012, 2013; Walker *et al*, 2013). It has been concluded that earlier diagnosis

could be achieved by instituting explicit criteria for investigation in primary care (National Institute for Health and Care Excellence (NICE), 2005; Scottish Intercollegiate Guidelines Network (SIGN), 2014). In the United Kingdom, the Be Clear on Cancer campaign funded by the Department of Health has been successful in urging individuals with particular symptoms to attend their GP surgery (Cancer Research UK, 2014a; Department of Health, 2012), and in selecting some of those individuals for investigation.

However, these studies have not established how much earlier diagnosis could be brought forward in patients presenting in primary care, nor how many individuals currently diagnosed at a

*Correspondence: Professor AE Ades; E-mail: t.ades@bristol.ac.uk

Received 4 August 2014; revised 24 September 2014; accepted 4 November 2014

© 2014 Cancer Research UK. All rights reserved 0007–0920/14

late stage could be diagnosed at earlier stages, when better treatment options are available.

The present study examines symptom-based investigation of lung and colorectal cancers from two points of view. First, it provides estimates of the Symptom Lead Time (SLT) distribution for these cancers. The SLT is the distribution of times between the presentation of symptoms *caused by cancer* in primary care and the eventual diagnosis of cancer (Lim *et al*, 2012; Ades *et al*, 2014). The SLT's role in a symptom-based programme is exactly analogous to that of the Lead Time in a standard screening programme. In an earlier, more technical, paper we estimated the SLT for lung cancer taking account of the fact that a proportion of diagnoses are already the result of GPs referring patients for chest X-ray (CXR) following presentation of symptoms in primary care. We showed that this had little impact on SLT. In this paper, we present analyses of SLT for a wide range of potential investigational criteria, in both lung and colorectal cancers.

A second objective of this paper is to estimate the proportion of cancers detected in a symptom-based programme in whom the symptoms that trigger the diagnostic investigation are actually caused by the cancer. This is a critical parameter, because cancers diagnosed as a result of symptoms that are *not* caused by cancer are being picked up serendipitously, by what is in effect by a random screening programme. It is therefore important to establish not only that those whose symptoms are caused by cancer enjoy a substantial lead time, and are therefore likely to benefit from earlier diagnosis, but also what proportion of all cancers detected by symptom-based programmes are serendipitous, and what the implications of this are for symptom-based investigative strategies.

In discussion, we examine the implications of our results for recent proposals in the United Kingdom to lower thresholds and widen access to diagnostic investigations for these and other cancers.

MATERIALS AND METHODS

This paper presents re-analyses of data collected on 247 lung cancer and 349 colorectal cancer cases diagnosed in Exeter (England) between 1998 and 2002, with five controls per case matched on age, sex, and GP practice (Hamilton *et al*, 2005; Hamilton, 2009). GP attendances over the 2-year period before cancer diagnosis were recorded, along with stage at diagnosis. Previous multivariable analyses had found that the following seven symptoms were associated with lung cancer: cough, dyspnoea, chest pain, fatigue, loss of weight, loss of appetite, and haemoptysis; and the following five with colorectal cancer: rectal bleeding, abdominal pain, loss of weight, diarrhoea, and constipation (Hamilton *et al*, 2005; Hamilton, 2009).

Sensitivity and specificity. We examined the sensitivity (proportion of cases meeting the symptom definition over the 2-year period), and specificity (proportion of controls meeting the definition) of: each individual symptom; any single symptom; any two symptoms occurring within 1 month, within 2, 3, 4, 5, and 6 months.

Among the potential symptom criteria with sensitivity over 40%, we picked out the following five for further study, as they represented a wide range of sensitivity and specificity. For lung these were (i) cough, (ii) any of the seven symptoms, (iii) any two of the seven symptoms within 3 months (including repeat presentations of the same symptom), (iv) any two different symptoms within 3 months, and (v) chest pain. For colorectal cancer: (i) rectal bleeding; (ii) any of the five symptoms; (iii) any two symptoms within 6 months; (iv) any two different symptoms within 6 months; and (v) abdominal pain. These criteria include the most sensitive criterion (any symptom), the criteria with the

highest diagnostic odds ratio (DOR), the criteria with the highest Youden Index, the most sensitive single symptom, and the most specific single symptom.

Symptom incidence and SLT distribution. We identified the number of times patients met the symptom definitions in each of eight 3-month intervals for lung cancer, and twelve 2-month intervals for colorectal cancer. For each diagnostic criterion, the rate of symptom presentation in cases and controls during each interval was estimated by Poisson regression. We interpreted the difference in incidence between cases and controls in each interval as the rate of symptoms caused by cancer. The SLT distribution is the incidence of symptoms caused by cancer in each interval, divided by the summed incidence of symptoms caused by cancer over the 2 years.

The proportion of symptoms in cases due to cancer. The incidence of symptoms caused by cancer, as defined above, divided by the total incidence in cases represents the proportion of symptoms in cases that are caused by cancer. We estimated this over the 2-year study period. We also estimated the proportion of symptoms caused by cancer that would be observed if a strategy of investigation based on these criteria were adopted in practice. This calculation takes into account the sojourn times of the cancers and allows for the possibility that symptoms caused by cancer could occur more than 2 years before diagnosis (see Discussion and Supplementary Materials).

Estimation was performed by Bayesian Markov chain Monte Carlo in WinBUGS (Lunn *et al*, 2000; Spiegelhalter *et al*, 2007). Further details of the statistical methods have been published previously (Ades *et al*, 2014).

RESULTS

Stage at diagnosis. Of the 247 lung cancer cases, the proportions in stages 1, 2, 3, 4 and unknown/missing are 12.6%, 6.9%, 12.6%, 22.3% and 45.7%, respectively. Of the 349 colorectal cancer cases, the proportions in Duke's A, B, C, D and unknown/missing are 13.2%, 35.8%, 28.7%, 10% and 12.3%, respectively.

Sensitivity and specificity of symptom criteria. The sensitivity and specificity of the different symptom criteria are shown in Table 1, along with the Youden index (sensitivity + specificity - 1), and DOR.

Symptom incidence. The overall incidence (% per year) of each of the investigative criteria in cases and controls is shown in Table 2. There are no systematic differences in symptom incidence between stages.

SLT distribution. The SLT for the criteria with the highest DOR is shown for each stage, and the same pattern is evident at all stages (Figure 1). Figure 2 gives the SLT averaged over all stages for each of the five lung and colorectal cancer investigative criteria. Despite the different sensitivity and specificity of the criteria, and their different incidence, the SLT distributions are virtually identical. The mean SLTs are between 4.7 and 6.0 months for lung cancer and between 4.1 and 5.0 months for colorectal cancer (Table 3). Median SLTs are between 2.4 and 3.2 for lung cancer, and between 2.0 and 2.7 for colorectal cancer.

Proportion of symptoms caused by cancer. The proportion of symptoms that are caused by cancer rises from between 12% and 38% to over 95% over the 2-year period (Figure 3). Averaged over the 2-year period of observation, between 66% and 83% of the occurrences of investigative criteria for lung cancer are due to lung cancer, and between 79% and 94% for colorectal cancer. Assuming a mean sojourn times of 4 years for both cancers (see Supplementary Materials and Discussion), the proportions that

Table 1. Sensitivity and specificity of the different symptom criteria

Investigative criteria	Sensitivity %	Specificity %	Youden index (95% CI)	DOR (95% CI)
Lung cancer				
Two different symptoms in 3 months	74.1	83.0	57.1 (51–63)	14.0 (10–19)
Cough	64.8	70.5	35.3 (29–49)	4.4 (3.3–5.9)
Any symptoms	91.9	52.8	44.7 (40–49)	12.7 (8.0–20)
Any two symptoms in 3 months	80.6	76.8	57.4 (52–63)	13.8 (9.8–19)
Chest pain	56.3	84.5	40.7 (34–47)	4.9 (3.6–6.7)
Colorectal cancer				
Any symptoms	86.2	73.2	59.5 (55–64)	17.2 (12–24)
Rectal bleeding	42.4	90.7	38.2 (33–44)	16.9 (12–23)
Any two symptoms in 6 months	60.2	89.4	49.6 (44–55)	12.7 (9.8–17)
Two different symptoms in 6 months	47.0	93.0	40.1 (35–46)	12.0 (9.1–16)
Abdominal pain	42.4	90.7	33.1 (28–38)	7.1 (5.5–9.3)

Abbreviations: CI = confidence interval; DOR = diagnostic odds ratio.

Table 2. Overall incidence (% per year) of each of the investigative criteria in cases and controls

Investigative criteria	Cases					Controls						
	Stage I (N = 31)	Stage II (N = 17)	Stage III (N = 31)	Stage IV (N = 55)	Missing/Unknown (N = 113)	Stage I (N = 155)	Stage II (N = 85)	Stage III (N = 155)	Stage IV (N = 275)	Missing/Unknown (N = 565)	Over all cases (N = 247)	Over all controls (N = 1235)
Lung cancer												
Two different symptoms in 3 months	55.0	65.3	60.0	60.0	70.8	15.8	14.1	11.3	10.9	12.8	64.7	12.7
Cough	56.4	62.4	63.2	69.0	66.4	26.4	22.4	25.8	22.7	22.9	65.0	23.6
Any symptoms	132.2	117.8	114.6	127.2	135.0	52.6	44.7	48.1	42.4	47.6	129.2	46.9
Any two symptoms in 3 months	68.0	77.1	74.4	75.5	85.9	22.3	19.4	17.1	15.1	18.1	79.3	17.9
Chest pain	24.5	39.4	23.2	32.8	33.2	8.1	7.6	7.7	5.8	9.8	31.2	8.3
Colorectal cancer												
Investigative criteria	Cases					Controls						
	Duke's A (N = 46)	Duke's B (N = 125)	Duke's C (N = 100)	Duke's D (N = 35)	Missing/Unknown (N = 43)	Duke's A (N = 230)	Duke's B (N = 624)	Duke's C (N = 500)	Duke's D (N = 175)	Missing/Unknown (N = 215)	Over all cases (N = 349)	Over all controls (N = 1744)
Any symptoms	87.0	85.2	80.0	95.7	79.1	19.1	18.2	17.9	20.6	18.1	84.3	18.5
Rectal bleeding	45.9	26.0	25.2	23.7	32.1	2.0	2.1	3.0	2.3	2.6	28.9	2.4
Any two symptoms in 6 months	47.9	46.4	43.6	57.4	44.5	5.0	6.3	6.3	6.3	8.4	46.7	6.4
Two different symptoms in 6 months	33.8	33.6	30.6	42.0	34.3	2.6	3.9	4.5	3.4	5.3	33.7	4.0
Abdominal pain	29.7	35.2	31.0	44.7	29.2	6.1	5.4	6.6	7.7	6.5	33.5	6.2

would be observed in practice are between 52% and 73% (lung), and between 68% and 88% (colorectal).

DISCUSSION

This study confirms results from much larger earlier studies (Hamilton *et al*, 2005; Hamilton, 2009; Hippisley-Cox and Coupland, 2011; Iyen-Omofoman *et al*, 2013), showing that lung and colorectal cancer cases present a range of characteristic symptoms in primary care at a substantially higher rate than matched controls, in the 2 years before diagnosis. However, these findings have previously been reported in the form of odds ratios and predictive values. Our analysis of SLT provides additional quantitative information about the lead time distribution that

might be achieved by various symptom-based investigation programmes. Specifically, we found that in both cancers the majority of the excess symptom presentation rate occurs in the last 3 months before diagnosis.

The choice of which symptom definitions are used as criteria for imaging appears to make little difference: the same result is obtained both for relatively sensitive but non-specific criteria, such as 'any symptom', and for more specific criteria, such as rectal bleeding for colorectal cancer. An important advantage of the present study is the information on stage at diagnosis. Symptom incidence (Table 2) and SLT results (Figure 1) show no relationship with stage. This suggests that few cases currently diagnosed at a late stage could be diagnosed at an earlier stage on the basis of symptoms caused by cancer presenting in primary care.

In an earlier more technical (Ades *et al*, 2014) paper on lung cancer, we explored the possibility that the short lead time was

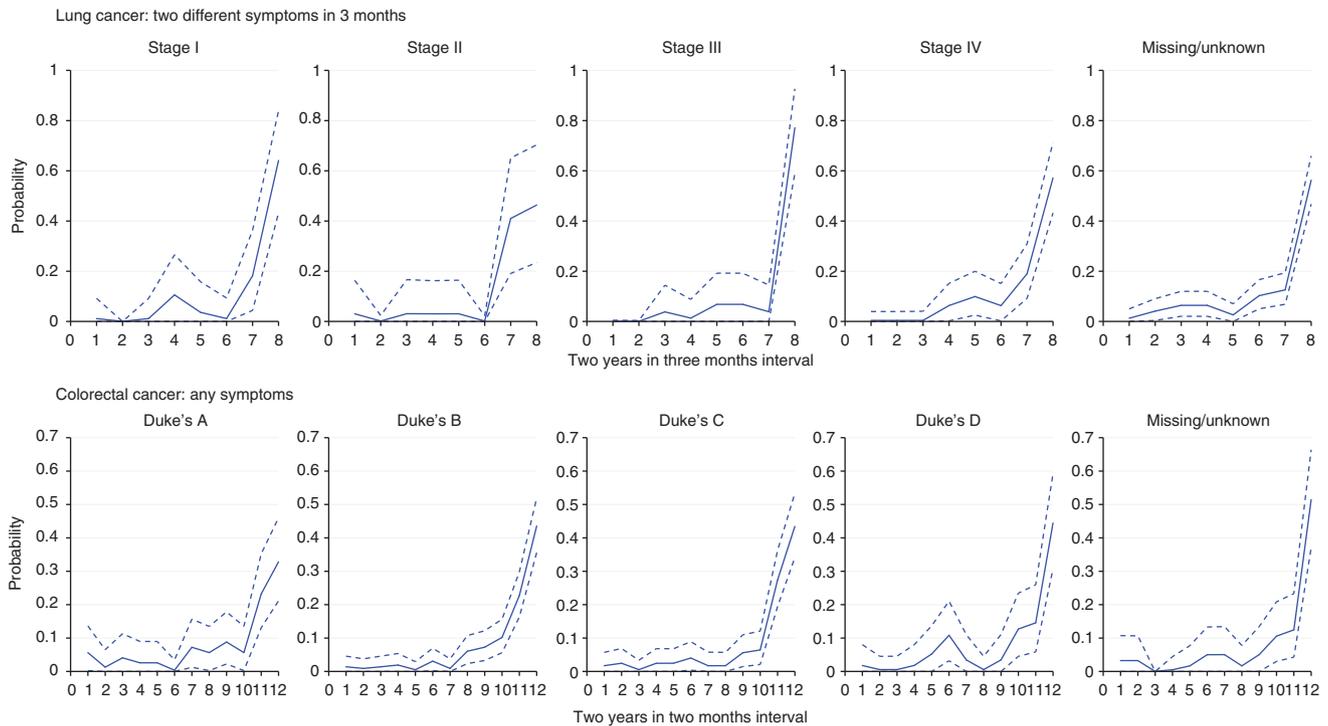


Figure 1. Symptom Lead Time distribution for ‘Any two symptoms in 3 months’ in lung cancer cases by stage and ‘Any symptoms’ in colorectal cases by stage.

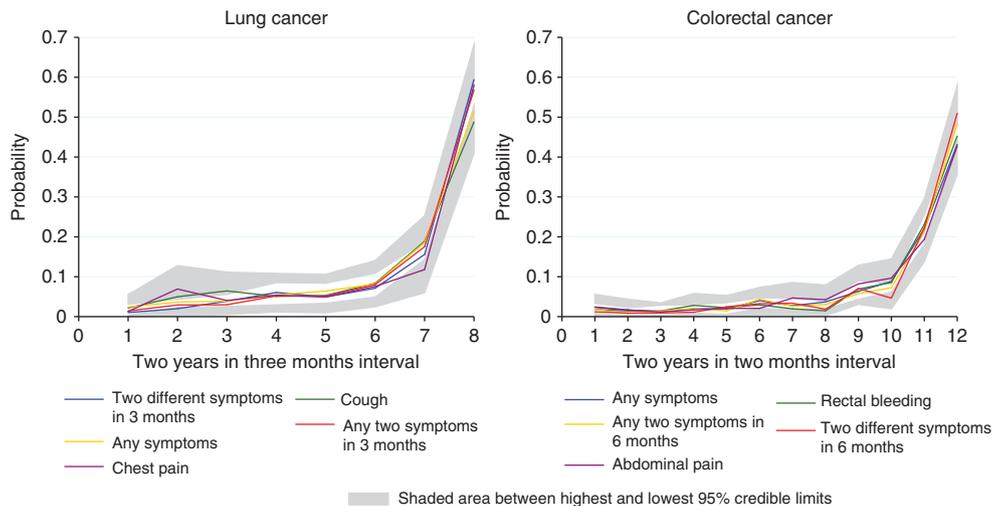


Figure 2. Symptom Lead Time distribution for five investigative criteria, lung and colorectal cancers. Shaded areas between the highest and lowest upper and lower 95% credible intervals.

partly the result of GP referral of patients with suggestive symptoms for chest X-ray. However, when patients with a CXR finding of ‘suspected lung cancer’ were excluded, the mean SLT was only 1 month longer. Thus, the estimates of SLT presented here may slightly underestimate the true SLT, but not to an extent that impacts on conclusions. Similar findings have been reported for ovarian cancer (Lim *et al*, 2012). As far as we know SLT for symptoms presenting in primary care has not been reported in any other cancers.

Potential limitations of our analysis should be recognised: the sample was small and restricted to the Exeter area, and whether it is representative of current practice can be questioned. However, there have been few significant changes in policy and practice since the data were collected in 1998–2002, and there no were special

organisational features that would have influenced speed or route of diagnosis, although increasing use of Computed Tomography (CT) scans may lead to earlier diagnoses. Although a proportion of lung cancer cases were unstaged, these tend to be those with advanced disease. If we assume that unstaged are all Stage IV, then the stage distribution in the data matches that of the two UK registries with the most complete data, 2004–2007 (Walters *et al*, 2013). Further, because stage appears to be unrelated to symptom incidence and SLT, it is unlikely that this missing data would affect conclusions.

Our main finding that median SLT is seldom more than 3 months suggests that most of those whose symptoms are caused by cancer might benefit relatively little in terms of time from symptom-based diagnosis. This does not necessarily mean that

the wide adoption of symptom-based diagnostic testing in primary care would have little effect on cancer outcomes. The avoidance of emergency presentation is a clear potential benefit, let alone any stage shift. However, it does raise questions about the rationale for such a strategy. To clarify this issue, we estimated the proportion of symptoms in cases whose symptoms were caused by cancer. For lung cancer, over the 24-month period before diagnosis, between 17% and 34% of symptoms presented by cases were *not* caused by the cancer, and between 6% and 21% for colorectal cancer. If these cases were referred for diagnostic imaging then *the detection of cancer would be essentially serendipitous*. Thus, individuals whose cancers are detected by a symptom-based programme are drawn from two distinct populations. Those whose symptoms were caused by cancer will benefit by having a lower risk of emergency admission, but, according to our SLT findings, might experience only relatively minor benefit in terms of stage at diagnosis as the programme offers a mean lead time of only 4–6 months.

The second group, whose cancer was detected serendipitously, have been diagnosed by what is, in effect, a screen carried at a random moment in time when they were asymptomatic with respect to their cancer. Mean lead times for these cases can be inferred from mean sojourn time estimates based on screening trials with multiple rounds (Day and Walter, 1984; Duffy *et al*, 1995; Prevost *et al*, 1998). The mean sojourn time for lung cancer is in the region of 4 years (Jang *et al*, 2013). For colorectal cancer a range of estimates between 2 and 6 years have been published

(Prevost *et al*, 1998; Brenner *et al*, 2011; Zheng and Rutter, 2012) with an average of ~ 4 years. Depending on the shape of the sojourn time distribution, the mean lead time from screening (assuming a perfect test) is usually greater than half the mean sojourn time, therefore about 2 years for both cancers. This is why standard screening programmes result in much earlier stage at diagnosis, better treatment options, and better cancer outcomes.

It has been implicitly assumed that symptom-based diagnosis identifies cases whose symptoms are caused by their cancer, and offers them diagnosis and treatment at an earlier stage. Instead, it seems that the benefits of a symptom-based cancer investigation strategy are experienced to a much higher extent than currently acknowledged by the cases whose symptoms are not caused by their cancer. It follows that the symptom-based programme that is based on the most sensitive and specific symptoms, although potentially more efficient in finding cancers, may offer the *least* clinical benefit to the cases that are detected.

This interpretation focusses attention on the results in Tables 1 and 4. The sensitivity, specificity, and proportion of symptoms due to cancer as estimated over the 2-year study period are not what we would expect to observe in a randomly chosen 2-year period. This is because the case-control design flatters the efficiency of the programme by over-sampling symptoms from the last 2 years of the sojourn time, when symptoms occur with a particularly high frequency. Therefore, we need to extrapolate Figure 3 back over the length of the sojourn time to obtain unbiased estimates. Assuming sojourn times of 4 years for both lung and colorectal cancers, we can derive projections for the proportion of symptoms caused by cancer that would be seen in practice. The estimates in Table 4 are conservative in that they allow for the possibility that cases present symptoms at a higher rate than controls for up to 3 years before diagnosis (see Supplementary materials for Materials and methods). For the symptom criteria with the highest DOR, the proportion of cases whose symptoms are caused by cancer is 73% for lung and 68% for colorectal. Note that less specific criteria have a higher rate of serendipitous detection: 'cough' would be caused by lung cancer in only 55% of cases.

These conclusions explain why the results of symptom-based programmes appear quite favourable, in spite of the low SLT. The Be Clear on Cancer campaign (Cancer Research UK, 2014a) has reported an increase in consultation rates (for persistent cough), increased chest X-ray referrals, an increased number of lung cancer diagnoses and an increased proportion of early stage diagnoses. However, *any* expansion of diagnostic testing necessarily generates

Table 3. Estimated mean and median Symptom Lead Time (SLT) in months for lung and colorectal cancers		
Investigative criteria	Mean SLT	Median SLT
Lung cancer		
Two different symptoms in 3 months	4.7 (4.0–5.5)	2.4 (2.0–2.8)
Cough	6.0 (4.9–7.1)	3.2 (2.6–5.1)
Any symptom	5.5 (4.8–6.3)	2.9 (2.6–3.9)
Any two symptoms in 3 months	4.9 (4.2–5.6)	2.6 (2.2–3.0)
Chest pain	5.5 (4.3–6.9)	2.5 (1.9–4.0)
Colorectal cancer		
Any symptoms	4.9 (4.3–5.6)	2.6 (2.1–3.3)
Rectal bleeding	4.6 (3.8–5.5)	2.4 (1.9–3.4)
Any two symptoms in 6 months	4.4 (3.8–5.1)	2.2 (1.8–2.9)
Two different symptoms in 6 months	4.1 (3.4–4.8)	2.0 (1.7–2.7)
Abdominal pain	5.0 (4.1–5.9)	2.7 (2.0–4.1)

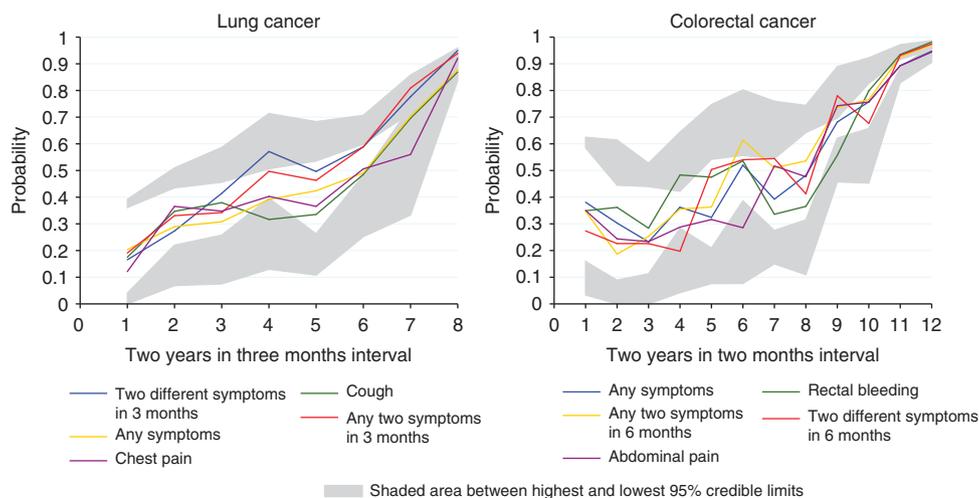


Figure 3. Probability of symptoms in cancer cases being due to cancer over 2-year period before diagnosis. Shaded areas between the most extreme upper and lower 95% credible intervals.

Table 4. Estimated percent of symptoms in lung and colorectal cases that are due to cancer, as observed percent during the 2-year study period, and predictions as to what would be observed in practice (posterior means and 95% credible interval (CrI))

Investigative criteria	During the 2-year period of observation (95% CrI)	In practice (95% CrI)
Lung cancer		
Two different symptoms in 3 months	83.1 (80.3–85.5)	73.3 (69.4–76.9)
Cough	67.0 (62.6–71.1)	54.6 (49.3–59.6)
Any symptoms	65.6 (62.3–68.8)	52.4 (48.4–56.3)
Any two symptoms in 3 months	80.2 (77.4–82.7)	69.2 (65.4–72.9)
Chest pain	77.9 (72.9–82.2)	67.4 (61.0–73.0)
Colorectal cancer		
Any symptoms	79.3 (77.0–81.5)	68.4 (65.2–71.4)
Rectal bleeding	93.6 (91.7–95.1)	88.4 (85.3–91.0)
Any two symptoms in 6 months	88.1 (86.0–90.0)	80.5 (77.3–83.4)
Two different symptoms in 6 months	90.1 (87.9–92.0)	83.1 (79.6–86.2)
Abdominal pain	84.1 (81.0–86.7)	74.5 (70.1–78.4)

more diagnoses, and therefore earlier diagnoses. Our findings suggest that these results should not be over-interpreted.

A further expansion of the National Awareness and Early Diagnosis Initiative (Cancer Research UK, 2014b) has been recently announced, involving lowering thresholds for diagnosis and increasing access to diagnostic services through the introduction of more routes to diagnosis, potentially including patient self-referral.

Symptom-based cancer investigation offers relatively little to the majority of patients with symptoms, those whose symptoms were caused by cancer, while offering a fair likelihood of early-stage diagnosis to those identified serendipitously. It offers nothing to the proportion of cancer sufferers who do not present with symptoms at all, or who present atypically. In addition, the general population with symptoms who do not have cancer must still suffer all the *disadvantages* of cancer screening programmes: false positives will undergo unnecessary, anxiety-provoking and invasive investigations; true positives are at risk of over-diagnosis leading to unnecessary major surgery. The Be Clear on Cancer campaign reported a 20% increase in diagnostic investigations for lung cancer, and a 9% increase in cancers detected, but also a 32% increase in urgent referrals for suspected lung cancer (Cancer Research UK, 2014a).

At the same time, the various safeguards that we associate with a planned screening programme, such as equal opportunity to benefit, logistic efficiency, and a national quality control programme, will *not* be available in a symptom-based strategy as currently conceived. Indeed, NHS England is actively encouraging local commissioning groups to adopt a service configuration of their own choice to promote earlier diagnosis. Among the options suggested are: a role for pharmacists in cancer referrals, lowering thresholds for urgent referral, and patient self-referral. We would urge that a rigorous examination of costs and benefits is carried out, perhaps by the National Screening Committee, before symptom-based investigation programmes are further expanded.

ACKNOWLEDGEMENTS

We 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. Ethics approval for the original study (CAPER study) was obtained from the North and East Devon research ethics committee. 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.

CONFLICT OF INTEREST

WH 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.

REFERENCES

- Ades A, Biswas M, Welton N, Hamilton W (2014) Symptom lead time distribution in lung cancer: natural history and prospects for early diagnosis. *Int J Epidemiol*; e-pub ahead of print 29 August 2014; doi:10.1093/ije/dyu174.
- Brenner H, Altenhofen L, Katalinic A, Lansdorp-Vogelaar I, Hoffmeister M (2011) Sojourn time of preclinical colorectal cancer by sex and age: estimates from the German national screening colonoscopy database. *Am J Epidemiol* **174**(10): 1140–1146.
- Cancer Research UK (2014a) *Be Clear on Cancer: Evaluation summary May 2014*. Cancer Research UK: London, UK.
- Cancer Research UK (2014b) National Awareness and Early Diagnosis Initiative (NAEDI). Available at <http://www.cancerresearchuk.org/cancer-info/spotcancerearly/naedi/> (accessed 15 July 2014).
- Day N, Walter S (1984) Simplified models of screening for chronic disease: Estimation procedures from mass screening programmes. *Biometrics* **40**(1): 1–14.
- Department of Health (2012) *Promoting early diagnosis of breast, bowel and lung cancers: first report 2010/11 local projects*. Department of Health: London, UK.
- Duffy S, Chen H, Tabar L, Day N (1995) Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. *Stat Med* **14**(14): 1531–1543.
- Friedman GD, Skilling JS, Udaltsova NV, Smith LH (2005) Early symptoms of ovarian cancer: a case-control study without recall bias. *Fam Pract* **22**(5): 548–553.
- Hamilton W (2009) The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. *Br J Cancer* **101**(Suppl 2): S80–S86.
- Hamilton W, Peters T, Round A, Sharp D (2005) What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. *Thorax* **60**(12): 1059–1065.
- Hippisley-Cox J, Coupland C (2011) Identifying patients with suspected lung cancer in primary care: Derivation and validation of an algorithm. *Br J Gen Pract* **61**(592): e715–e723.
- Iyen-Omofoman B, Tata L, Baldwin D, Smith C, Hubbard R (2013) Using socio-demographic and early clinical features in general practice to identify people with lung cancer earlier. *Thorax* **68**(5): 451–459.
- Jang H, Kim S, Wu D (2013) Bayesian lead time estimation for the Johns Hopkins Lung Project data. *J Epidemiol Global Health* **3**(3): 157–163.
- Lim AWW, Mesher D, Gentry-Maharaj A, Balogun N, Jacobs I, Menon U, Sasieni P (2012) Predictive value of symptoms for ovarian cancer: comparison of symptoms reported by questionnaire, interview, and general practitioner notes. *J Natl Cancer Inst* **104**: 1–11.
- Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* **10**: 325–337.

- National Institute for Health and Care Excellence (NICE) (2005) *Referral Guidelines for suspected cancer*. NICE: London, UK.
- Prevost TC, Launoy G, Duffy SW, Chen HH (1998) Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. *Am J Epidemiol* **148**(6): 609–619.
- Scottish Intercollegiate Guidelines Network (SIGN) (2014) *SIGN 137: Management of Lung Cancer*. SIGN: Edinburgh, UK.
- Shephard E, Stapley S, Neal R, Rose P, Walter F, Hamilton W (2012) Clinical features of bladder cancer in primary care. *Br J Gen Pract* **62**: e598–e604.
- Shephard E, Stapley S, Neal R, Rose P, Walter F, Hamilton W (2013) Clinical features of kidney cancer in primary care. *Br J Gen Pract* **63**: e250–e255.
- Spiegelhalter D, Thomas A, Best N, Lunn D (2007) WinBUGS User Manual Version 1.4 January 2003. Upgraded to Version 1.4.3. Available from: <http://www.mrc-bsu.cam.ac.uk/bugs/> (accessed 2014).
- Stapley S, Peters T, Neal R, Rose P, Walter F, Hamilton W (2012) The risk of pancreatic cancer in symptomatic patients in primary care: a large case-control study using electronic records. *Br J Cancer* **106**(12): 1940–1944.
- Stapley S, Peters T, Neal R, Rose P, Walter F, Hamilton W (2013) The risk of oesophago-gastric cancer in symptomatic patients in primary care: a large case-control study using electronic records. *Br J Cancer* **108**(1): 25–31.
- Walker S, Hyde C, Hamilton W (2013) Risk of uterine cancer in symptomatic women in primary care: case-control study using electronic records. *Br J Gen Pract* **63**: e643–e648.
- Walters S, Maringe C, Coleman M, Peake M, Butler J, Young N, Bergstrom S, Hanna L, Jakobsen E, Kolbeck K, Sundstrom S, Engholm G, Gavin A, Gjerstorff M, Hatcher J, Johannesen T, Linklater K, McGahan C, Steward J, Tracey E, Turner D, Richards M, Rachet B, IMW G (2013) Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: A population-based study, 2004–2007. *Thorax* **68**(6): 551–564.
- Zheng W, Rutter C (2012) Estimated mean sojourn time associated with hemocult SENSE for detection of proximal and distal colorectal cancer. *Cancer Epidemiol Biomarkers Prev* **21**(10): 1722–1730.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)