

Keywords: lung cancer; referral; diagnosis; treatment; interval; target; time

Factors associated with timeliness of post-primary care referral, diagnosis and treatment for lung cancer: population-based, data-linkage study

L F Forrest^{*,1,2}, J Adams^{1,2}, M White^{1,2} and G Rubin^{1,3}

¹Fuse, UKCRC Centre for Translational Research in Public Health, Newcastle University, Newcastle upon Tyne NE2 4AX, UK;

²Institute of Health & Society, Newcastle University, Baddiley Clark Building, Richardson Road, Newcastle upon Tyne NE2 4AX, UK and ³Wolfson Research Institute, Durham University, Queen's Campus, Stockton on Tees TS17 6BH, UK

Background: The NHS Cancer Plan for England set waiting time targets for cancer referral (14 days from GP referral to first hospital appointment) and treatment (31 days from diagnosis, 62 days from urgent GP referral). Interim diagnostic intervals can also be calculated. The factors that influence timely post-primary care referral, diagnosis and treatment for lung cancer are not known.

Methods: Northern and Yorkshire Cancer Registry, Hospital Episode Statistics and lung cancer audit data sets were linked. Logistic regression was used to investigate the factors (socioeconomic position, age, sex, histology, co-morbidity, year of diagnosis, stage and performance status (PS)) that may influence the likelihood of referral, diagnosis and treatment within target, for 28 733 lung cancer patients diagnosed in 2006–2010.

Results: Late-stage, poor PS and small-cell histology were associated with a higher likelihood of post-primary care referral, diagnosis and treatment within target. Older patients were significantly less likely to receive treatment within the 31-day (odds ratio (OR) = 0.79, 95% confidence interval (CI) 0.69–0.91) and 62-day target (OR = 0.80, 95% CI 0.67–0.95) compared with younger patients.

Conclusions: Older patients waited longer for treatment and this may be unjustified. Patients who appeared ill were referred, diagnosed and treated more quickly and this 'sicker quicker' effect may cancel out system socioeconomic inequalities that might result in longer time intervals for more deprived patients.

Early diagnosis of lung cancer is thought to be important for improving outcomes. Survival is better for patients diagnosed at an early stage because they are more likely to be suitable for receipt of curative treatment (Richards, 2009). Delays between the onset of cancer symptoms and receipt of treatment could result in potentially resectable tumours becoming inoperable, and this may contribute to the poorer survival of UK cancer patients compared with that found in other European countries (Richards, 2009).

In the United Kingdom, there are a number of different routes to diagnosis. Three of these diagnostic pathways reflect the urgency of the referral (emergency presentation at hospital or via the GP, urgent 2-week GP referral and other GP referral; Elliss-Brookes *et al*, 2012). Theoretical models of the pathway from first symptom to cancer treatment identify key diagnostic (including referral) and treatment intervals, and related healthcare system (primary and secondary care) settings (Hansen *et al*, 2008; Walter *et al*, 2012).

*Correspondence: Dr LF Forrest; E-mail: Lynne.Forrest2@ncl.ac.uk

Received 20 March 2014; revised 14 July 2014; accepted 24 July 2014; published online 9 September 2014

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

An early model of cancer delay, the Anderson model, attributed the majority of delay to patient factors (Andersen *et al*, 1995). The model has been updated to also consider tumour and healthcare system factors (Walter *et al*, 2012), as these are likely to make important contributions to the time to diagnosis and treatment.

Three intervals are currently the subject of performance management within the NHS in England. Since 2000, urgent referrals for suspected cancer have been required to have a first hospital appointment (FHA) within 14 days from the date of GP referral (referral interval). Since 2005, target intervals of 62 days from the date of urgent GP referral to first treatment and 31 days from diagnosis (decision-to-treat) to first treatment (treatment intervals) have been in place (Department of Health, 2000). Interim diagnostic interval targets of 31 days from GP referral and 17 days from FHA can be inferred (Figure 1).

Although research has explored factors that might influence the time spent within pathway intervals for breast and colorectal cancer (Ramirez *et al*, 1999; Mitchell *et al*, 2008), there is little definitive evidence on the factors that are important for lung cancer. In a review of potential reasons for delay in lung cancer care (14 studies), variables including stage, co-morbidity, atypical symptoms, income, age, sex, rural residence and distance to healthcare were examined (Olsson *et al*, 2009). The authors were unable to draw any clear conclusions as they reported that the quality of included studies was poor. The majority of studies considered only single factors and did not take into account potential confounders such as age, stage, histology and co-morbidity (Olsson *et al*, 2009).

In this study, cancer registry, Hospital Episode Statistics (HES) and lung cancer audit (LUCADA) data sets were linked to investigate the factors (socioeconomic position (SEP), age, sex, histology, co-morbidity, year of diagnosis, stage and performance status (PS)) that may influence the likelihood of post-primary care referral, diagnosis and treatment within target times.

MATERIALS AND METHODS

Data sources and linkage. We analysed a linked data set reported previously (Forrest *et al*, 2014c). Data for 29 385 patients with a primary diagnosis of lung cancer (ICD10 C33 and C34), diagnosed between 1 January 2006 and 31 December 2010, were obtained from the Northern and Yorkshire Cancer Registry and Information Centre (NYCRIS, 2012). Of these, 652 had tumour registration based on death certification only and so were excluded from analyses, leaving an eligible cohort of 28 733.

Data on SEP, age, sex, histology, year of diagnosis, GP referral date, FHA date, diagnosis date, treatment dates and details of receipt of treatment (surgery, chemotherapy and radiotherapy) were obtained from the registry data. Co-morbidity data were obtained from HES and used to calculate a co-morbidity score. Data on stage and PS were obtained from the national LUCADA, a non-mandatory register of clinical information on patients diagnosed with lung cancer comprising a subset of registry patients (66% nationally in 2006, increasing to 93% in 2010; NHS Information Centre, 2012).

Records were allocated a unique, randomly generated, key number, derived from the NHS number by NYCRIS. Data from the three data sources were anonymised and supplied by NYCRIS. The HES and LUCADA data were then linked to the registry data using this key.

Variables of interest. SEP was measured using the income domain of the Index of Multiple Deprivation (IMD; HM Government, 2013), grouped into quintiles for analysis. This is an area-based measure of SEP, where Q5 is the most deprived and Q1 the least deprived. Age at diagnosis was categorised into four groups for analysis: <60, 60–69, 70–79 and 80+ years. Year of diagnosis was included to examine system changes over time.

Histology was classified as confirmed non-small-cell lung cancer (NSCLC), including adenocarcinoma, large-cell carcinoma, non-small-cell carcinoma and squamous cell carcinoma subtypes; small-cell lung cancer (SCLC); and other histology (including unspecified carcinoma, neoplasm, other specified carcinomas and carcinoid tumours).

A weighted co-morbidity score was calculated by NYCRIS using the Charlson co-morbidity index (CCI; Charlson *et al*, 1987) a validated instrument used previously to determine co-morbidity in cancer patients. No HES-linked data were available for patients diagnosed in 2009–2010 as, due to national problems in calculating the co-morbidity score, there was a time lag in data availability. Patients without a CCI score were split into those who did and did not have a HES linkage. CCI score was categorised as 0, 1–2, 3+, missing (HES-linked data available but no CCI score recorded) or unavailable (no HES-linked data available).

Stage was assigned using the TNM staging system and categorised as I, II, III, IV or missing (Sobin and Wittekind, 1997). PS is a measure of general well-being for cancer patients, assessed on a scale of 0 (asymptomatic) to 4 (bedridden) using the Eastern Co-operative Group PS scale (NICE, 2005) and was categorised as 0, 1–2, 3–4 or missing. Stage and PS data were entered into LUCADA by trusts participating in the audit.

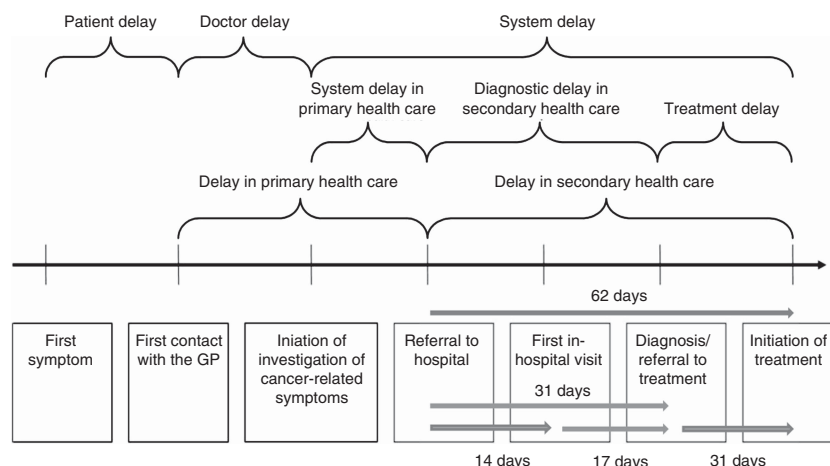


Figure 1. Intervals and target times on the cancer diagnosis and treatment pathway (adapted from Hansen *et al*, 2008).

Table 1 shows the referral and treatment intervals included in this study (Forrest *et al*, 2014b), alongside target times for these. The referral interval was categorised as: GP referral date to FHA ≤ 14 days (within target), GP referral > 14 days and no GP referral. Interim diagnostic intervals and the target times that can be inferred are also shown in Table 1. These were calculated as follows: GP referral to diagnosis ($62 - 31 = 31$ days) and FHA to diagnosis ($31 - 14 = 17$ days).

Between 2006 and 2009, date of diagnosis was determined by cancer registries as the first time the tumour was identified either by imaging or histology in the case notes. In 2010, UK Association of Cancer Registry guidance was followed. The guidance uses a priority hierarchy with six categories to determine date of diagnosis; the date of histological or cytological confirmation of malignancy (which can be derived from three different hierarchical time points); date of hospital admission; and down to, with lowest priority, date of death if cancer is only identified at autopsy (Weller *et al*, 2012).

Analysis. Median referral, diagnostic and treatment times, with inter-quartile range (IQR), were calculated for each time interval. The number and percentage of patients referred, diagnosed and treated within target were also calculated for each time interval, by each included variable. Differences in the likelihood of referral, diagnosis and treatment within target by SEP, age, sex, histology, year of diagnosis, co-morbidity score, stage and PS were explored using logistic regression models mutually adjusted for all co-variables. As the aim was to explore which factors might be important for time spent within each interval, the variable selection procedure was based on an underlying conceptual framework rather than on formal stepwise methods (Kirkwood and Sterne, 2003). Variables were selected for inclusion if they had previously been identified as potentially important. Referral interval was adjusted for in the models for the FHA to diagnosis and diagnosis to treatment intervals. Type of first treatment was included in the models for treatment intervals. Receipt of treatment was included in the models for diagnostic intervals.

Patients were eligible for inclusion within an interval if they had a date recorded for the interval start and end point. Records were excluded from analyses if they had a negative interval for any time period, for example, if they had a GP referral date later than diagnosis date, as this was likely to be a data entry error. Records were included if they had 0 time between dates (for example, treatment date was recorded on the same day as date of diagnosis). To include these records in the analysis, 0 time intervals were recoded as 0.1. Analysis was also restricted to those cases with interval dates within 1 year of the previous interval end point.

Odds ratios (ORs) with 95% confidence intervals (CIs) for the likelihood of receipt of referral, diagnosis and treatment within target for each variable were reported. Likelihood ratio tests were used to determine the overall significance of each categorical

variable. Analysis was carried out in Stata v12.0 (StataCorp, College Station, TX, USA).

RESULTS

Data and descriptive statistics. Of the full cohort (28 733), 7769 (27.0%) had stage and 8885 (30.9%) had a PS score recorded in LUCADA, and 18 650 (64.9%) had a linked HES record, of whom 8475 (29.5%) had a CCI score recorded. All 28 733 patients had a date of diagnosis, 28 704 (99.9%) had FHA date recorded, 15 452 (53.8%) had a GP referral date and 15 373 (53.5%) received any treatment within 1 year of diagnosis. Once patients with negative intervals and intervals longer than 1 year were excluded, 14 507 patients had GP referral and FHA dates recorded, and also GP referral and diagnosis dates; 27 555 had FHA and diagnosis dates; 14 692 had diagnosis and treatment dates; and 9542 had GP referral and treatment dates. The demographic and clinical characteristics of the patients included in each analysis are shown in columns 1–4 of Tables 2–4.

Referral interval: likelihood of FHA within 14 days from GP referral. Median time from GP referral to FHA was 10 days (IQR 6–17), $n = 14 507$. Of those referred by their GP, 70.4% had FHA within the 14-day target from GP referral. The data set included both urgent and non-urgent referrals.

Greater deprivation was associated with a lower likelihood of FHA within 14 days of GP referral in the multivariable analysis (OR = 0.83, CI 0.73–0.94), but the lowest likelihood was for those in the middle SEP groups. No simple pattern by age emerged. Late-stage cancer, SCLC, poor PS and referral post-2006 were associated with increased likelihood of FHA within 14 days of referral, but sex and CCI score were not (Table 2).

Diagnostic interval: likelihood of diagnosis within 31 days from GP referral. In those patients with GP referral and diagnosis dates ($n = 14 507$), median time from GP referral to diagnosis was 13 days (IQR 7–24) and 82.6% were diagnosed within the 31-day target.

Neither SEP nor age was associated with likelihood of diagnosis within 31 days of GP referral overall, although those in the middle SEP groups had significantly lower likelihood of diagnosis within 31 days (Table 3). Patients with poorer PS, later-stage cancer and those with SCLC had an increased likelihood of diagnosis within 31 days of GP referral. Those diagnosed in 2007–2009 had a higher likelihood of diagnosis within 31 days.

Diagnostic interval: likelihood of diagnosis within 17 days from FHA. In those patients with FHA and diagnosis dates ($n = 27 555$), median time from FHA to diagnosis was 0 days (IQR 0–0) and 91.4% of patients were diagnosed within 17 days.

Neither SEP nor age was associated with likelihood of diagnosis within 17 days of FHA (Table 3). Patients with poorer PS, later-stage cancer and those with SCLC had an increased likelihood of timely diagnosis. Those diagnosed in 2007–2009 had a higher likelihood of timely diagnosis.

Treatment interval: likelihood of treatment within 31 days from diagnosis. Of the 14 692 patients who were treated within 1 year, median time to treatment was 35 days (IQR 21–55) and 43.1% (6333) were treated within the 31-day target time from diagnosis (Table 4).

There was no association between SEP and likelihood of treatment within 31 days of diagnosis ($P = 0.41$). Patients diagnosed with SCLC were significantly more likely to receive treatment within target compared with those diagnosed with NSCLC (OR = 4.53, CI 4.06–5.04). Older patients were significantly less likely to receive treatment within the 31-day target (OR = 0.79, CI 0.69–0.91).

Table 1. Referral, diagnostic and treatment intervals

Interval	Definition of interval	Target time (days)
Referral	GP referral date to first hospital appointment date	14
Diagnostic	GP referral date to diagnosis date ^a	31 ^b
	First hospital appointment date to diagnosis date ^a	17 ^b
Treatment	Diagnosis date to first treatment date	31
	GP referral date to first treatment date	62

Abbreviation: GP = general practitioner.
^aInterim target intervals.
^bCalculated interim target times.

Table 2. Post-primary care referral interval: likelihood of FHA within 14 days of GP referral^a

	GP referral & FHA date	FHA within 14 days of GP referral		Multivariable analysis (n = 14 507, R ² = 7.09) ^b			
	N	N	%	OR	95% CI		P-value
IMD	14 507	10 218	70.4				0.0004
1 (least deprived)	1709	1263	73.9	1.00			
2	2028	1430	70.5	0.85	0.73	0.99	
3	2456	1714	69.8	0.79	0.68	0.91	
4	3447	2363	68.6	0.74	0.65	0.85	
5 (most deprived)	4867	3448	70.8	0.83	0.73	0.94	
Age range	14 507	10 218	70.4				0.019
< 60	2008	1450	72.2	1.00			
60–69	4143	2975	71.8	0.97	0.86	1.10	
70–79	5263	3628	68.9	0.86	0.76	0.97	
80 +	3093	2165	70.0	0.89	0.77	1.02	
Sex	14 507	10 218	70.4				0.60
Female	6601	4681	70.9	1.00			
Male	7906	5537	70.0	1.02	0.95	1.10	
Histology	14 507	10 218	70.4				<0.001
NSCLC	8803	6077	69.0	1.00			
SCLC	1898	1479	77.9	1.64	1.45	1.85	
Other	3806	2662	69.9	1.03	0.94	1.13	
Diagnosis year	14 507	10 218	70.4				<0.001
2006	2465	1235	50.1	1.00			
2007	2575	1475	57.3	1.34	1.20	1.50	
2008	3062	2479	81.0	4.26	3.77	4.81	
2009	3156	2501	79.3	4.39	3.79	5.09	
2010	3249	2528	77.8	3.94	3.37	4.60	
CCI score	14 507	10 218	70.4				<0.001
0	1899	1279	67.4	1.00			
1–2	1389	936	67.4	0.89	0.76	1.04	
3 +	294	201	68.4	0.80	0.60	1.05	
Missing	5384	3560	66.1	1.24	1.10	1.40	
Unavailable	5541	4242	76.6	0.96	0.83	1.10	
Stage	14 507	10 218	70.4				<0.001
I	764	504	66.0	1.00			
II	393	282	71.8	1.32	1.00	1.73	
III	1558	1169	75.0	1.49	1.22	1.81	
IV	2307	1848	80.1	1.77	1.47	2.15	
Missing	9485	6415	67.6	1.45	1.19	1.76	
Performance status	14 507	10 218	70.4				0.001
0	1351	990	73.3	1.00			
1–2	3198	2394	74.9	1.02	0.88	1.19	
3–4	1217	988	81.2	1.36	1.11	1.67	
Missing	8741	5846	66.9	0.94	0.79	1.12	

Abbreviations: CCI score=Charlson co-morbidity score; CI=confidence interval; FHA=first hospital appointment; GP=general practitioner; IMD=Index of Multiple Deprivation; NSCLC=non-small-cell lung cancer; OR=odds ratio; SCLC=small-cell lung cancer.

^aExcluding those with FHA >1 year from GP referral, diagnosis >1 year from FHA, any negative time interval.

^bMutually adjusted for socioeconomic position, age, sex, histology, year of diagnosis, co-morbidity score, stage and PS.

Those receiving surgery were less likely to receive treatment within target times compared with those receiving chemotherapy. Those with poor PS and later-stage cancer were more likely to receive treatment within target.

Treatment interval: likelihood of treatment within 62 days from GP referral. Of the 9542 patients who were referred by their GP and had treatment, median time from GP referral to first treatment date was 56 days (IQR 39–79) and 61.2% of those who received treatment were treated within 62 days (Table 4), although we had

no details of how many of these were urgent referrals. An overall association between SEP and likelihood of starting treatment within the target time was found ($P=0.04$), with those in the middle SEP groups significantly less likely to start treatment within 62 days.

Early-stage patients, older patients and those receiving surgery were significantly less likely to start treatment within 62 days, whereas late-stage patients, those with poor PS and those receiving chemotherapy, were more likely to do so. The likelihood of receiving treatment within guidelines significantly improved over time.

Table 3. Diagnostic interval: likelihood of diagnosis within 31 days from GP referral, and of diagnosis within 17 days from FHA^a

	GP referral to diagnosis							FHA to diagnosis							
	GP referral & diagnosis date	Diagnosis within 31 days of GP referral		Multivariable analysis (n = 14 507, R ² = 4.56) ^b			FHA & diagnosis date	Diagnosis within 17 days of FHA		Multivariable analysis (n = 27 555, R ² = 8.67) ^c					
		N	N	%	OR	95% CI		P-value	N	N	%	OR	95% CI		P-value
IMD	14 507	11 978	82.6				0.08	27 555	25 197	91.4					0.65
1 (least deprived)	1709	1436	84.0	1.00				3257	2990	91.8	1.00				
2	2028	1669	82.3	0.90	0.75	1.07		4020	3666	91.2	0.91	0.77	1.09		
3	2456	2006	81.7	0.84	0.71	1.00		4658	4239	91.0	0.88	0.75	1.04		
4	3447	2820	81.8	0.83	0.71	0.97		6437	5887	91.5	0.90	0.77	1.05		
5 (most deprived)	4867	4047	83.2	0.94	0.80	1.09		9183	8415	91.6	0.93	0.80	1.08		
Age range	14 507	11 978	82.6				0.43	27 555	25 197	91.4					0.89
< 60	2008	1661	82.7	1.00				3503	3166	90.4	1.00				
60–69	4143	3437	83.0	1.06	0.91	1.22		7262	6562	90.4	1.01	0.88	1.17		
70–79	5263	4314	82.0	0.96	0.84	1.11		9830	8989	91.4	1.01	0.88	1.16		
80 +	3093	2566	83.0	0.97	0.83	1.15		6960	6480	93.1	0.96	0.82	1.13		
Sex	14 507	11 978	82.6				0.009	27 555	25 197	91.4					0.03
Female	6601	5410	82.0	1.00				12 695	11 571	91.2	1.00				
Male	7906	6568	83.1	1.13	1.03	1.23		14 860	13 626	91.7	1.11	1.01	1.21		
Histology	14 507	11 978	82.6				<0.001	27 555	25 197	91.4					<0.001
NSCLC	8803	7134	81.0	1.00				14 484	12 915	89.2	1.00				
SCLC	1898	1674	88.2	1.77	1.52	2.07		3350	3114	93.0	1.63	1.41	1.89		
Other	3806	3170	83.3	1.06	0.94	1.19		9721	9168	94.3	1.45	1.29	1.64		
Diagnosis year	14 507	11 978	82.6				<0.001	27 555	25 197	91.4					<0.001
2006	2465	1895	76.9	1.00				5440	5082	93.4	1.00				
2007	2575	2139	83.1	1.49	1.29	1.71		5619	5298	94.3	1.16	0.99	1.35		
2008	3062	2706	88.4	2.25	1.94	2.61		5541	5223	94.3	1.14	0.97	1.34		
2009	3156	2754	87.3	2.41	2.03	2.87		5489	5111	93.1	1.13	0.95	1.36		
2010	3249	2484	76.5	1.05	0.89	1.26		5466	4483	82.0	0.34	0.29	0.41		
CCI score	14 507	11 978	82.6				<0.001	27 555	25 197	91.4					<0.001
0	1899	1527	80.4	1.00				3841	3492	90.9	1.00				
1–2	1389	1102	79.3	0.86	0.72	1.02		3377	3082	91.3	0.94	0.80	1.12		
3 +	294	223	75.9	0.68	0.50	0.92		886	798	90.1	0.84	0.65	1.08		
Missing	5384	4591	85.3	1.53	1.32	1.77		9897	9367	94.6	1.54	1.33	1.80		
Unavailable	5541	4535	81.8	1.10	0.94	1.29		9554	8458	88.5	1.16	1.00	1.35		
Stage	14 507	11 978	82.6				<0.001	27 555	25 197	91.4					<0.001
I	764	539	70.6	1.00				1116	879	78.8	1.00				
II	393	304	77.4	1.52	1.14	2.03		526	434	82.5	1.46	1.10	1.92		
III	1558	1313	84.3	2.10	1.70	2.60		2172	1942	89.4	2.28	1.85	2.81		
IV	2307	2053	89.0	2.91	2.36	3.60		3583	3340	93.2	3.30	2.68	4.06		
Missing	9485	7769	81.9	1.63	1.32	2.02		20 158	18 602	92.3	1.71	1.39	2.11		
Performance status	14 507	11 978	82.6				0.0007	27 555	25 197	91.4					0.002
0	1351	1069	79.1	1.00				1759	1481	84.2	1.00				
1–2	3198	2675	83.7	1.22	1.03	1.44		4631	4107	88.7	1.17	0.99	1.39		
3–4	1217	1085	89.2	1.66	1.30	2.12		2077	1952	94.0	1.44	1.13	1.83		
Missing	8741	7149	81.8	1.18	0.97	1.44		19 088	17 657	92.5	1.43	1.17	1.73		
Any treatment ^d	14 507	11 978	82.6				<0.001	27 555	25 197	91.4					<0.001
No	4974	4241	85.3	1.00				12 863	12 228	95.1	1.00				
Yes	9533	7737	81.2	0.76	0.67	0.85		14 692	12 969	88.3	0.45	0.40	0.51		
GP referral date	14 507	11 978	82.6					27 555	25 197	91.4					0.79
No								13 048	12 105	92.8	1.00				
Yes								14 507	13 092	90.3	0.99	0.90	1.09		

Abbreviations: CCI score=Charlson co-morbidity score; CI=confidence interval; FHA=first hospital appointment; GP=general practitioner; IMD=Index of Multiple Deprivation; NSCLC=non-small-cell lung cancer; OR=odds ratio; SCLC=small-cell lung cancer.

^aExcluding those with FHA > 1 year from GP referral, diagnosis > 1 year from FHA, any negative time interval.

^bMutually adjusted for socioeconomic position (SEP), age, sex, histology, year of diagnosis, co-morbidity score, stage and PS.

^cMutually adjusted for SEP, age, sex, histology, year of diagnosis, co-morbidity score, stage, PS and GP referral date.

^dWithin 1 year of diagnosis.

Table 4. Treatment interval: likelihood of treatment within 31 days from diagnosis, and of treatment within 62 days from GP referral^a

	Diagnosis to treatment							GP referral to treatment							
	Diagnosis & treatment date	Treatment within 31 days of diagnosis		Multivariable analysis (n = 14 692, R ² = 9.64) ^b			GP referral & treatment date	Treatment within 62 days of GP referral		Multivariable analysis (n = 9542, R ² = 9.18) ^c					
		N	N	%	OR	95% CI		P-value	N	N	%	OR	95% CI		P-value
IMD	14 692	6333	43.1				0.41	9542	5839	61.2					0.04
1 (least deprived)	1844	822	44.6	1.00				1186	760	64.1	1.00				
2	2207	961	43.5	0.97	0.84	1.10		1381	850	61.6	0.89	0.75	1.06		
3	2493	1060	42.5	0.89	0.78	1.02		1611	969	60.2	0.83	0.70	0.97		
4	3353	1457	43.5	0.94	0.83	1.07		2186	1310	59.9	0.79	0.68	0.92		
5 (most deprived)	4795	2033	42.4	0.91	0.81	1.03		3178	1950	61.4	0.88	0.75	1.01		
Age range	14 692	6333	43.1				<0.001	9542	5839	61.2					<0.001
< 60	2795	1356	48.5	1.00				1749	1172	67.0	1.00				
60–69	5071	2277	44.9	0.90	0.81	0.99		3313	2105	63.5	0.93	0.81	1.05		
70–79	5145	2052	39.9	0.75	0.68	0.83		3355	1913	57.0	0.73	0.64	0.83		
80+	1681	648	38.6	0.79	0.69	0.91		1125	649	57.7	0.80	0.67	0.95		
Sex	14 692	6333	43.1				0.19	9542	5839	61.2					0.11
Female	6611	2958	44.7	1.00				4260	2592	60.9	1.00				
Male	8081	3375	41.8	0.95	0.89	1.02		5282	3247	61.5	1.07	0.98	1.17		
Histology	14 692	6333	43.1				<0.001	9542	5839	61.2					<0.001
NSCLC	10 170	3647	35.9	1.00				6864	3947	57.5	1.00				
SCLC	2499	1863	74.6	4.53	4.06	5.04		1554	1313	84.5	3.05	2.61	3.57		
Other	2023	823	40.7	1.27	1.14	1.41		1124	579	51.5	0.87	0.76	0.99		
Diagnosis year	14 692	6333	43.1				<0.001	9542	5839	61.2					0.14
2006	2910	1294	44.5	1.00				1659	965	58.2	1.00				
2007	2960	1239	41.9	0.89	0.80	0.99		1734	1020	58.8	1.02	0.88	1.18		
2008	2870	1035	36.1	0.65	0.58	0.73		1967	1213	61.7	1.10	0.96	1.27		
2009	3001	1133	37.8	0.71	0.62	0.83		2040	1258	61.7	1.17	0.97	1.41		
2010	2951	1632	55.3	1.60	1.37	1.88		2142	1383	64.6	1.28	1.05	1.57		
CCI score	14 692	6333	43.1				0.45	9542	5839	61.2					<0.001
0	2211	914	41.3	1.00				1315	776	59.0	1.00				
1–2	1537	580	37.7	0.91	0.79	1.05		828	426	51.5	0.77	0.64	0.93		
3+	330	137	41.5	1.07	0.83	1.37		147	74	50.3	0.72	0.50	1.03		
Missing	5843	2434	41.7	1.01	0.90	1.13		3834	2363	61.6	1.12	0.97	1.30		
Unavailable	4771	2268	47.5	1.03	0.90	1.17		3418	2200	64.4	1.11	0.94	1.32		
Stage	14 692	6333	43.1				<0.001	9542	5839	61.2					<0.001
I	870	218	25.1	1.00				626	233	37.2	1.00				
II	384	94	24.5	0.86	0.65	1.15		314	136	43.3	1.09	0.82	1.45		
III	1529	567	37.1	1.36	1.11	1.66		1176	756	64.3	1.61	1.29	2.00		
IV	1909	1024	53.6	2.52	2.07	3.07		1366	1067	78.1	2.87	2.29	3.60		
Missing	10 000	4430	44.3	1.82	1.49	2.22		6060	3647	60.2	1.65	1.32	2.07		
Performance status	14 692	6333	43.1				0.0001	9542	5839	61.2					0.02
0	1599	630	39.4	1.00				1247	750	60.1	1.00				
1–2	3417	1502	44.0	1.11	0.97	1.27		2476	1636	66.1	1.08	0.93	1.27		
3–4	447	247	55.3	1.66	1.31	2.10		296	212	71.6	1.31	0.96	1.77		
Missing	9229	3954	42.8	1.00	0.85	1.16		5523	3241	58.7	0.88	0.73	1.06		
First treatment ^d	14 692	6333	43.1				<0.001	9542	5839	61.2					<0.001
Chemotherapy	6587	3470	52.7	1.00				4486	3273	73.0	1.00				
Surgery	2687	864	32.2	0.76	0.68	0.85		1748	654	37.4	0.37	0.32	0.42		
Radiotherapy	5418	1999	36.9	0.77	0.70	0.84		3308	1912	57.8	0.73	0.66	0.82		
GP referral	14 692	6333	43.1				<0.001								
No GP referral date	5159	2354	45.6	1.00											
FHA ≤ 14 days	6683	2933	43.9	0.92	0.85	1.00									
FHA > 14 days	2850	1046	36.7	0.74	0.67	0.82									

Abbreviations: CCI score = Charlson co-morbidity score; CI = confidence interval; FHA = first hospital appointment; GP = general practitioner; IMD = Index of Multiple Deprivation; NSCLC = non-small-cell lung cancer; OR = odds ratio; SCLC = small-cell lung cancer.

^aExcluding those with FHA > 1 year from GP referral, diagnosis > 1 year from FHA, any negative time interval.

^bMultivariable adjusted for socioeconomic position (SEP), age, sex, histology, year of diagnosis, co-morbidity score, stage, PS and GP referral.

^cMultivariable adjusted for SEP, age, sex, histology, year of diagnosis, co-morbidity score, stage and PS.

^dWithin 1 year of diagnosis.

DISCUSSION

Principal findings. To our knowledge, this is the first study to examine factors associated with referral, diagnosis and treatment within target for lung cancer in England. Generally, those with SCLC, those with poorer PS, patients diagnosed in more recent years, and those with more advanced stage cancer were more likely to have post-primary care referral, diagnosis and treatment within target. Older patients were less likely to receive first treatment within the target time. Patients undergoing surgery and radiotherapy as their first treatment were less likely to receive treatment within target than those receiving chemotherapy as a first treatment.

Socioeconomic inequalities in the time from GP referral to FHA, diagnosis and treatment were found, but were not present for the interim intervals from FHA to diagnosis or from diagnosis to treatment. This would suggest that SEP inequalities originate from the GP referral to FHA interval. However, a linear trend was not seen. Patients in the middle SEP groups were less likely to receive FHA, diagnosis or treatment within target, compared with the least deprived SEP group, but this was not consistently observed for the most deprived SEP group.

Strengths and weaknesses. The use of multiple data set linkage is a strength of this study. There are good data completeness and validity within the registry data set. Linkage to HES and LUCADA data allowed inclusion of a wide range of factors that are potentially important for post-primary care referral, diagnosis and treatment within target for lung cancer. Previous studies examining lung cancer delay mainly conducted only univariable analyses and did not include potential confounders such as age, stage, histology and co-morbidity (Olsson *et al*, 2009). We were able to take all these factors, as well as others such as PS and SEP, into account. However, a lack of concordance in recording of data between registry and LUCADA data sets has been found (Riaz *et al*, 2010).

The type of referral route may influence time to diagnosis and treatment. Nearly half of the patients (46%) did not have a GP referral date recorded, but it was not possible to determine from our data how many of these patients had been referred via a different route, how many had a GP referral but the date was not recorded or which were urgent GP referrals. A large study examining routes to diagnosis also found that low numbers (41%) of lung cancer patients were referred via their GP (as urgent or non-urgent referrals; Elliss-Brookes *et al*, 2012). Those who did have a GP date had longer time to treatment than those who did not. This reflects previous findings (Allgar and Neal, 2005) and suggests that those without a GP referral date may be presenting as emergency cases and thus receive more urgent investigation and treatment.

We examined the interim diagnostic interval from FHA to diagnosis. However, the FHA and diagnosis date often appeared to be recorded as the same date, which may be due to variability in the recording of date of diagnosis in the registry data set. A number of different dates can be used to determine date of diagnosis. Inconsistent application of the rules could affect calculation of the diagnostic and treatment intervals and thus introduce error, although there is no evidence to suggest bias in recording. It is possible that, for those without a GP referral date, the FHA may be an emergency hospital visit where diagnosis is made on that day. Another UK study also found that the median time from FHA to diagnosis was 0 days for lung cancer and for five other cancer types (Allgar and Neal, 2005) using patient survey data. As the same result was found using two different data sources, this suggests that many patients do receive their diagnosis at the FHA.

The high levels of missing data for stage, PS and co-morbidity were a weakness of the study and the inclusion of 'missing' categories for these variables may result in bias. Multiple imputation was not possible as over 50% of the data were missing for these variables (White *et al*, 2011). However, although a high proportion of missing data may make absolute estimates of proportions unreliable, it is unlikely to matter substantially for estimation of the independent effects of each variable, within each fully adjusted model. As the majority of other studies did not conduct multivariable analysis (Olsson *et al*, 2009), there is little other research to compare our results to, in order to assess whether missing data could account for any of the observed patterns. SEP was measured at an area-based level using IMD, which is a further limitation, as area-based measures may not accurately reflect individual-level circumstances.

Interpretation of results and comparison with other studies

Diagnostic interval. Diagnostic delay has been implicated as a factor that contributes to the poorer survival of UK cancer patients compared with the European average (Richards, 2009). However, the impact of delay on lung cancer survival is unclear (Jensen *et al*, 2002) and other studies suggest that those who experience shorter system delay may, in fact, have poorer lung cancer survival as they are at the 'sicker' end of the spectrum (Myrdal *et al*, 2004; Forrest *et al*, 2014a), with symptoms that result in more urgent referral and diagnosis.

Two studies from Sweden (Berglund *et al*, 2010) and Denmark (Dalton *et al*, 2011), countries with similar healthcare systems to the United Kingdom, found socioeconomic inequalities in time from referral to diagnosis, with higher SEP patients having more prompt diagnosis. We found socioeconomic inequalities from GP referral to diagnosis and, within this interval, from GP referral to FHA, but not from FHA to diagnosis. This replicates previous UK findings (Neal and Allgar, 2005). The evidence therefore suggests that socioeconomic diagnostic inequalities are found in the GP referral to FHA interval rather than the secondary care FHA to diagnostic interval.

Referral interval. It has been suggested that socioeconomic differences in communicating and presenting cancer symptoms to health professionals may result in longer delays for those who are less 'convincing' (Corner and Brindle, 2011). Lower SEP patients may be less articulate, have poorer health literacy and be more culturally distant from doctors, whereas more educated cancer patients are better able to describe symptoms and thus speed up the referral process (Hansen *et al*, 2008; Forrest, 2013). However, a nonlinear trend in socioeconomic inequalities in the likelihood of being referred to secondary care within time interval guidelines was found. Those in the middle SEP groups were less likely to receive timely referral than those in the least deprived SEP group, but this was not seen for those in the most deprived SEP group.

The waiting time paradox suggests that sicker people are referred more quickly as they are more obviously ill, but also less likely to survive (Torrington *et al*, 2011), and so results in a positive association bias. Our findings of an inverse relationship between PS and likelihood of referral within target support this theory. This may be a possible explanation as to why the most deprived patients were not significantly less likely to receive a timely referral or diagnosis compared with the most affluent, as we might have expected. Patients with lower SEP had poorer PS (results not shown) and so were likely to be in generally poorer health. The results of this 'sicker quicker' effect may, therefore, act to effectively 'cancel out' any potential system inequalities that might result in longer referral and diagnosis intervals for more deprived patients.

In a previous study, age and sex did not appear to influence the likelihood of urgent referral for lung cancer (Allgar *et al*, 2006). We also found that sex was not associated with likelihood of timely referral and no clear pattern of referral was found by age.

Patients subsequently diagnosed with later-stage cancer and those with poor PS had a higher likelihood of being referred within 14 days. Stage, PS and co-morbidity score can all be considered as proxy measures for how ill a patient is, which may influence referral behaviour. Urgent referral requires the presence of 'alarm' symptoms that are more likely to be present as the disease advances.

A previous small study found that co-morbidity delayed diagnosis in around 23% of lung cancer patients seen by GPs (Bjerager *et al*, 2006). Symptoms (for example cough, weight loss or breathlessness) may be ascribed to a known co-morbidity (for example, COPD) rather than lung cancer. However, no clear-cut pattern was observed in referral or diagnosis time by co-morbidity score in this study, but those with poorer PS had increased likelihood of referral and diagnosis within target.

Both CCI score and PS can be used as proxy measures of general well-being and it is difficult to say how well they capture this (Forrest *et al*, 2014b). CCI is a validated instrument for measuring co-morbidity (Charlson *et al*, 1987) but it is a fairly crude measure and, as it only contains details of conditions requiring in-patient care, it is likely to underestimate co-morbidity (Forrest *et al*, 2014c). It may be that the number of co-morbidities is not a particularly sensitive measure of patient well-being and PS may be a better marker.

Treatment interval. SEP was not associated with treatment within target from diagnosis. However, in the referral to treatment interval, those in the middle SEP groups appeared less likely to receive treatment within the 62-day guidelines, compared with the least deprived group. Those with SCLC, later-stage cancer and poorer PS were more likely to receive timely treatment within target. Again the 'sicker quicker' phenomenon may explain this.

The previous evidence for time spent on the care pathway by age is mixed, with a UK study using patient survey data finding that younger patients had longer time intervals (Neal and Allgar, 2005). In agreement with a Swedish study using medical record data, which found that older patients had longer time to lung cancer treatment (Lovgren *et al*, 2008), we also found that older patients were significantly less likely to receive treatment within target time. This was seen even when factors such as stage, co-morbidity and PS were taken into account, which might indicate that older patients are experiencing unwarranted delays.

Implications for policy and practice, and further research. Due to changes in the way that date of diagnosis was calculated in 2010, it was difficult to be sure that any differences seen in this year were system changes as a result of the introduction of targets or just due to the change in rules for determining date of diagnosis. Consistent determination of GP referral, FHA, diagnosis and treatment dates and robust systems for reliably capturing this data are necessary for accurate calculation of referral, diagnosis and treatment intervals.

However, as improvements over time were found pre-2010, it does appear that introduction of NICE guideline targets (NICE, 2005) has resulted in a greater proportion of patients being referred, diagnosed and treated within target times.

In this study, patients who underwent surgery as a first treatment were less likely to be treated within target compared with patients undergoing chemotherapy. It has been suggested that those who receive curative surgery may undergo a longer period of preliminary investigation with a greater number of diagnostic and staging procedures (Devbhandari *et al*, 2007). However, there may be capacity limitations on the pathway to surgery and future

research should investigate whether it is possible to improve the timeliness of surgical treatment.

Older lung cancer patients were less likely to be treated within target times. Ageism in cancer care has been highlighted as an important problem that needs to be addressed (National Cancer Equality Initiative, 2012; Lawler *et al*, 2014) and our results add to the evidence for this. Clinicians need to ensure that older patients do not undergo unnecessary and unjustified delays in treatment. Research into the reasons for later treatment in older patients is required.

The patterns of results found in this study need to be clarified in other data sets, ideally with lower levels of missing stage, PS and co-morbidity data. Our findings suggest that more deprived patients are likely to appear more ill (as measured by PS) and the 'sicker quicker' effect may neutralise any patient and system inequalities that might result in longer referral, diagnostic and treatment intervals for these patients, potentially relating to poorer communication and health literacy. Better evidence of inequalities in health literacy by SEP, and the role this may have in delay, is required. Investigation into patient and primary care delay before GP referral, and of inequalities within these intervals, is required.

CONCLUSIONS

In this cohort, the main factor determining timely post-primary care referral, diagnosis and treatment for lung cancer was how sick patients appeared. Improvements in meeting NICE time interval targets were seen over time. However, older patients, who were more likely to have poor PS, were less likely to receive treatment within target, indicating possible age discrimination. It is possible that the 'sicker quicker' effect may effectively 'cancel out' system inequalities that might result in longer time intervals for more deprived patients.

ACKNOWLEDGEMENTS

LFF was funded by the ESRC (PhD studentship ES/I020926/1) as a member of Fuse, the Centre for Translational Research in Public Health (www.fuse.ac.uk). MW is director, GR a senior investigator and JA is funded as a staff member of Fuse. Fuse is a UK Clinical Research Collaboration (UKCRC) Public Health Research Centre of Excellence. Funding for Fuse from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the National Institute for Health Research, under the auspices of the UKCRC, is gratefully acknowledged. The views expressed in this paper do not necessarily represent those of the funders or UKCRC. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We would like to thank staff at NYCRI for extracting and supplying the data for linkage. Ethical Approval was applied for through the Integrated Research Application System (IRAS) for NHS Research Ethics Committee (REC) approval. A favourable ethical opinion was obtained from the Proportionate Review sub-committee of the NRES Committee East of England REC on the 13 December 2011 (REC reference 11/EE/0537).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

LFF designed the study, obtained the data, conducted the analysis and wrote the first draft of the manuscript. JA, MW and GR had the initial idea for the study, obtained the funding and were involved in data interpretation and critical revision of the manuscript. The final manuscript was approved by all authors before submission. All authors will act as guarantors.

REFERENCES

- Allgar VL, Neal RD (2005) Delays in the diagnosis of six cancers: analysis of data from the National Survey of NHS Patients: Cancer. *Br J Cancer* **92**(11): 1959–1970.
- Allgar VL, Neal RD, Ali N, Leese B, Heywood P, Proctor G, Evans J (2006) Urgent GP referrals for suspected lung, colorectal, prostate and ovarian cancer. *Br J Gen Pract* **56**: 355–362.
- Andersen BL, Cacioppo JT, Roberts DC (1995) Delay in seeking a cancer diagnosis: delay stages and psychophysiological comparison processes. *Br J Soc Psychol* **34**: 33–52.
- Berglund A, Holmberg L, Tishelman C, Wagenius G, Eaker S, Lambe M (2010) Social inequalities in non-small cell lung cancer management and survival: a population-based study in central Sweden. *Thorax* **65**(4): 327–333.
- Bjerager M, Palshof T, Dahl R, Vedsted P, Olesen F (2006) Delay in diagnosis of lung cancer in general practice. *Br J Gen Pract* **56**: 863–868.
- Charlson M, Pompei P, Ales K, MacKenzie C (1987) A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* **40**: 373–383.
- Corner J, Brindle L (2011) The influence of social processes on the timing of cancer diagnosis: a research agenda. *J Epidemiol Community Health* **65**: 477–482.
- Dalton SO, Frederiksen BL, Jacobsen E, Steding-Jessen M, Osterlind K, Schuz J, Osler M, Johansen C (2011) Socioeconomic position, stage of lung cancer and time between referral and diagnosis in Denmark, 2001–2008. *Br J Cancer* **105**(7): 1042–1048.
- Department of Health (2000) The NHS cancer plan: a plan for investment, a plan for reform. Available at http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4009609 (last accessed 28 February 2013).
- Devbhndari M, Soon S, Quennell P, Barber P, Krysiak P, Shah R, Jones M (2007) UK waiting time targets in lung cancer treatment: are they achievable? Results of a prospective tracking study. *J Cardiothorac Surg* **2**(1): 5.
- Elliss-Brookes L, McPhail S, Ives A, Greenslade M, Shelton J, Hiom S, Richards M (2012) Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets. *Br J Cancer* **107**(8): 1220–1226.
- Forrest LF (2013) Why are socioeconomic inequalities in receipt of treatment found for lung cancer? *Lung Cancer Manage* **2**(3): 177–180.
- Forrest LF, Adams J, Rubin G, White M (2014a) The role of receipt and timeliness of treatment in socioeconomic inequalities in lung cancer survival: population-based, data linkage study. *Thorax*. e-pub ahead of print 12 June 2014; doi:10.1136/thoraxjnl-2014-205517.
- Forrest LF, Sowden S, Rubin G, White M, Adams J (2014b) Socio-economic inequalities in patient, primary care, referral, diagnostic, and treatment intervals on the lung cancer care pathway: Protocol for a systematic review and meta-analysis. *Syst Rev* **3**: 30.
- Forrest LF, White M, Rubin G, Adams J (2014c) The role of patient, tumour and system factors on socioeconomic inequalities in lung cancer treatment: population-based study. *Br J Cancer* **111**(3): 608–618.
- Hansen R, Olesen F, Sorensen H, Sokolowski I, Sondergaard J (2008) Socioeconomic patient characteristics predict delay in cancer diagnosis: a Danish cohort study. *BMC Health Serv Res* **8**(1): 49.
- HM Government (2013) English Indices of Deprivation 2010. <http://data.gov.uk/dataset/index-of-multiple-deprivation>.
- Jensen AR, Mainz J, Overgaard J (2002) Impact of delay on diagnosis and treatment of primary lung cancer. *Acta Oncol* **41**(2): 147–152.
- Kirkwood BR, Sterne JA (2003) Regression modelling. In *Essential Medical Statistics*. 2nd edn pp 315–342. Blackwell Publishing Ltd: Oxford, UK.
- Lawler M, Selby P, Aapro M, Duffy S (2014) Ageism in cancer care. *BMJ* **348**: g1614.
- Lovgren M, Levealahti H, Tishelman C, Runesdotter S, Hamberg K (2008) Time spans from first symptom to treatment in patients with lung cancer—the influence of symptoms and demographic characteristics. *Acta Oncol* **47**(3): 397–405.
- Mitchell E, Macdonald S, Campbell NC, Weller D, Macleod U (2008) Influences on pre-hospital delay in the diagnosis of colorectal cancer: a systematic review. *Br J Cancer* **98**(1): 60–70.
- Myrdal G, Lambe M, Hillerdal G, Lamberg K, Agustsson T, Stahle E (2004) Effect of delays on prognosis in patients with non-small cell lung cancer. *Thorax* **59**(1): 45–49.
- National Cancer Equality Initiative (2012) The impact of patient age on clinical decision-making in oncology. Available at <https://www.gov.uk/government/publications/the-impact-of-patient-age-on-clinical-decision-making-in-oncology>.
- Neal RD, Allgar VL (2005) Sociodemographic factors and delays in the diagnosis of six cancers: analysis of data from the "National Survey of NHS patients: Cancer". *Br J Cancer* **92**: 1971–1975.
- NHS Information Centre (2012) *National Lung Cancer Audit Report 2012*. NHS Information Centre for Health and Social Care: Leeds, UK. Available at www.hqip.org.uk/.../Lung-Cancer-National-Audit-Report-pub-2012.pdf.
- NICE (2005) *Clinical Guideline 24. Lung Cancer: the diagnosis and treatment of lung cancer*. NICE: London, UK.
- NYCRIS (2012) Northern and Yorkshire Cancer Registry and Information Centre. Available at <http://www.nycris.nhs.uk/about/> (last accessed 2 August 2013).
- Olsson JK, Schultz EM, Gould MK (2009) Timeliness of care in patients with lung cancer: a systematic review. *Thorax* **64**(9): 749–756.
- Ramirez AJ, Westcombe AM, Burgess CC, Sutton S, Littlejohns P, Richards MA (1999) Factors predicting delayed presentation of symptomatic breast cancer: a systematic review. *Lancet* **353**(9159): 1127–1131.
- Riaz SP, Linklater K, Horton M, Peake M, Moller H, Luchtenborg M (2010) Lung cancer data in the National Cancer Data Repository, Hospital Episode Statistics and National Lung Cancer Audit datasets. Available at www.ncin.org.uk/view?rid=1624.
- Richards MA (2009) The National Awareness and Early Diagnosis Initiative in England: assembling the evidence. *Br J Cancer* **101**(S2): S1–S4.
- Sobin LH, Wittekind CH (1997) *TNM Classification of Malignant Tumors*. 5th edn (John Wiley & Sons: New York, NY, USA).
- Torring ML, Frydenberg M, Hansen RP, Olesen F, Hamilton W, Vedsted P (2011) Time to diagnosis and mortality in colorectal cancer: a cohort study in primary care. *Br J Cancer* **104**(6): 934–940.
- Walter F, Webster A, Scott S, Emery J (2012) The Andersen Model of Total Patient Delay: a systematic review of its application in cancer diagnosis. *J Health Serv Res Policy* **17**(2): 110–118.
- Weller D, Vedsted P, Rubin G, Walter FM, Emery J, Scott S, Campbell C, Anderson RS, Hamilton W, Olesen F, Rose P, Nafees S, van Rijswijk E, Hiom S, Muth C, Beyer M, Neal RD (2012) The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. *Br J Cancer* **106**: 1262–1267.
- White IR, Royston P, Wood AM (2011) Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* **30**(4): 377–399.

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.