Comparing primary and secondary health-care use between diagnostic routes before a colorectal cancer diagnosis: Cohort study using linked data

J R Sheringham*,[1,2], T Georghiou[1], X A Chitnis[3] and M Bardsley[1]

[1]Nuffield Trust, 59 New Cavendish Street, London W1G 7LP, UK; [2]University College London, 1-19 Torrington Place, London WC1E 7HB, UK and [3]The Royal Marsden NHS Foundation Trust, Downs Road, Sutton SM2 5PT, UK

Background: Survival in cancer patients diagnosed following emergency presentations is poorer than those diagnosed through other routes. To identify points for intervention to improve survival, a better understanding of patients' primary and secondary health-care use before diagnosis is needed. Our aim was to compare colorectal cancer patients' health-care use by diagnostic route.

Methods: Cohort study of colorectal cancers using linked primary and secondary care and cancer registry data (2009–2011) from four London boroughs. The prevalence of all and relevant GP consultations and rates of primary and secondary care use up to 21 months before diagnosis were compared across diagnostic routes (emergency, GP-referred and consultant/other).

Results: The data set comprised 943 colorectal cancers with 24% diagnosed through emergency routes. Most (84%) emergency patients saw their GP 6 months before diagnosis but their symptom profile was distinct; fewer had symptoms meeting urgent referral criteria than GP-referred patients. Compared with GP-referred, emergency patients used primary care less (IRR: 0.85 (95% CI 0.78–0.93)) and urgent care more frequently (IRR: 1.56 (95% CI 1.12; 2.17)).

Conclusions: Distinct patterns of health-care use in patients diagnosed through emergency routes were identified in this cohort. Such analyses using linked data can inform strategies for improving early diagnosis of colorectal cancer.

In England, more than a quarter of patients with colorectal cancer are diagnosed as an emergency presentation, that is, following a visit to Accident and Emergency (A&E) or an emergency admission to hospital (Elliss-Brookes *et al*, 2012). Short term survival in these cases is poor when compared with other routes to diagnosis even when age and case mix are taken into account (Downing *et al*, 2013; McPhail *et al*, 2013). This has led to the interpretation of emergency presentations as an indicator of preventable diagnostic delay in colorectal cancer. As such, there could be scope to improve survival in colorectal cancer by reducing their prevalence (Hamilton, 2012). Although patient characteristics associated with emergency diagnosis are relatively

well characterised, the factors that lead to an emergency presentation and the extent to which they are tractable are less well understood.

McPhail *et al* (2013) propose that an understanding of how patients use primary and secondary care before their diagnosis is needed to develop effective strategies to improving cancer survival by reducing emergency presentation. Strategies have been developed to promote GP recognition of symptoms (Hamilton *et al*, 2013) and greater access to diagnostics has shown some effect on reducing diagnostic intervals (Neal *et al*, 2014). However, cancers in those diagnosed as emergencies may manifest in different ways to those diagnosed through GP-referred routes (Cleary *et al*, 2007).

In this case, initiatives to promote access to diagnostics for common symptoms and increased recognition of these symptoms in GPs and patients could have limited impact on the prevalence of emergency presentations.

In addition, wider national health-care policy to manage urgent care use (reducing A&E visits and emergency hospital admissions) may also affect emergency presentations in cancer (NHS England, 2013). This policy has led to campaigns to discourage patients from visiting A&E and encourage primary care use. If some patients diagnosed with cancer following an emergency presentation habitually access health-care through A&E rather than visiting a GP, then a generic policy to reduce A&E use may also reduce emergency presentations in cancer.

Many records-based studies have been limited to examining just part of the cancer diagnostic pathway in either secondary or primary care. However, with the prospect of computerised person-level records across the care pathway, (Sheather and Brannan, 2013) new methods for auditing care through secure data linkage are emerging (Dixon and Bardsley, 2012). In this study, we used linked, routinely-collected, primary and secondary care records supplemented with cancer registry data from outer North East London to identify targets for investigation or intervention to improve the diagnostic pathway in colorectal cancer. Specifically, we sought to address three objectives:

1. Prevalence and characteristics of primary care use in the year before diagnosis by route to diagnosis;
2. Variations in rates of urgent and scheduled health-care use by route to diagnosis up to 21 months prior to diagnosis, to capture not just activity connected with a cancer diagnosis but longer term service use patterns also;
3. Patient and health service factors associated with emergency presentations.

MATERIALS AND METHODS

We examined anonymised person-level records to study health service use up to 21 months before diagnosis for all colorectal cancer cases that were diagnosed between 2009–2011 in four London boroughs.

Population. Redbridge, Barking and Dagenham, Havering and Waltham Forest serve a diverse population of ∼1 million comprising 188 practices. The boroughs vary widely in their age profile and socioeconomic circumstances, with Havering having double the proportion of people over 65 years as Waltham Forest (20 and 10%, respectively). For the purposes of this paper, the boroughs have been anonymised.

The data set. Public Health England Knowledge and Intelligence (London) supplied Cancer Registry data on all cases of colorectal cancer (i.e., with ICD10 codes C18-C20) for individuals resident in the four outer North East London boroughs 2009–2012. The Registry data set includes a set of variables recording stage of cancer at diagnosis. The Registry also has collated information from a number of sources to generate a date of diagnosis using consistent criteria across all cases.

We obtained a primary and secondary care data set from the four clinical commissioning groups (CCGs) for all people in the locally registered populations with at least one recorded diagnosis of colorectal cancer (selected using the ICD10 codes above in any inpatient diagnostic field) during 2007–2012.

The data set comprised all inpatient, outpatient and A&E episodes (derived from Secondary Uses Service data) and all primary care records (from local administrative systems) in this same period. These data included information on patients'

socio-demographic characteristics (age, sex, lower super output area (LSOA) of residence) in addition to the dates of all primary care episodes (with corresponding Read Codes) and all hospital episodes.

To assign an area-based indicator of socioeconomic deprivation, LSOA was linked to national Indices of Multiple Deprivation 2010 scores and ranks and divided into quintiles (McLennan et al, 2011). To ensure identifiable data did not leave the NHS, the CCGs and Public Health England pseudonymised patient identifiers using a unique key before linkage at Nuffield Trust. We then excluded cases of colorectal cancer where there was evidence of any prior cancer in Cancer Registry data.

We compared the age and stage of unmatched records (those with a diagnosis in the Cancer registry data who were not recorded in the CCG data) to those we were able to match, to examine the extent to which the cases included in the cohort and those excluded were different.

Generating variables. We assigned a route to diagnosis to all cancer cases by adapting the method developed by Elliss-Brookes et al (2012). We first examined all episodes within 28 days of the date of diagnosis to find the relevant hospital episode (outpatient, inpatient, both unscheduled and scheduled, or A&E visits) closest to the diagnostic date. Second, we looked back at patients' secondary care activity in the 6 months prior to this episode. In contrast to Elliss-Brookes et al (2012), we sought to focus only on health service related to the diagnosis of cancer. Therefore, we excluded episodes of hospital activity 28 days to 6 months before diagnosis if there was no record of colonoscopy, major or minor resection, or inpatient or outpatient treatment in general surgery, colorectal surgery or gastroenterology.

We then sought to identify a source of referral to hospital. In most cases, there were just 1–2 hospital episodes remaining in the 28 days to 6 months before the diagnosis date, and we examined referral data for the first episode. In a small number of cases, there were more than three hospital episodes in the 6 months before diagnosis ($n = 20$). In these cases, we selected the first hospital episode with a primary care referral.

Finally, we used this information to allocate patients to four distinct routes:

1. Emergency presentation: patients with a record of A&E activity only, or where inpatient or outpatient activity resulted from a referral from A&E, or where referral from GP was an emergency;
2. GP urgent/2-week wait: patients diagnosed as a result of an initial outpatient episode where the priority of referral was recorded as urgent or as a 2-week wait;
3. GP routine/unknown: patients diagnosed as a result of an initial outpatient episode where the priority of referral was recorded as routine and those diagnosed through an inpatient episode only, where it was not possible to distinguish priority of referral;
4. Consultant, other, unknown: patients diagnosed as a result of a referral from a consultant or other health-care professional, self-referral, screening or where it was not possible to identify a route. There were no data in the registry data set on whether cancers were detected by screening, so it was not possible to isolate a distinct 'screening' route or exclude these cases from analysis.

We merged the two GP-referred routes in order to increase power and because there were no significant differences between results with aggregated GP routes and separate ones. We present data, therefore, using just three routes: Emergency, GP and Consultant/other/unknown. In analysis of emergency presentations, we use 'Emergency' vs 'Other' (i.e., GP and Consultant/Other/Unknown combined) to focus on the distinct characteristics of emergency routes to diagnosis.

We generated an aggregated stage variable that used either Tumour, Node, Metastases or Dukes stage to minimise the impact of missing data in our analysis, thereby creating three groups: 'Early' (Dukes A or B or no node involvement), 'Late' (record of Dukes stage C or D, node involvement or metastases) or missing (insufficient data to assign a stage). We collapsed this to a binary variable: early *vs* late/not known because a higher proportion of cancer staging information has been found to be missing in patients that die rapidly after diagnosis (Downing *et al*, 2013).

We generated variables to capture whether patients discussed any signs or symptoms with their GP in the year before diagnosis that might plausibly trigger a referral or investigation of colorectal cancer (called 'relevant' symptoms throughout the rest of this paper). We used Read Codes as a proxy for the presence of these 'relevant' symptoms, in common with other studies that have used these data to determine a diagnostic interval (Tate *et al*, 2009; Neal *et al*, 2014). The Read Code list generated (for details, see supplementary data) was as wide as possible to take account of the fact that Read Codes are not used uniformly between or within practices, nor are they intended to capture mention of every symptom. It encompassed codes for symptoms, examinations, diagnoses, investigations and prescriptions related to bowel-specific complaints (e.g., rectal bleeding, altered bowel habit, loose stools, diarrhoea and constipation), other symptoms that might trigger a bowel investigation (e.g., abdominal pain, and swelling) and investigation or presence of non-specific symptoms that might trigger cancer investigation (anaemia, fatigue and weight loss). It was based on codes used in other studies (Marshall *et al*, 2011; Taylor and Radford, 2012), plus additional codes added following examination of our data and using the Read Code hierarchy.

Analysis. We first describe the characteristics of our sample with exclusions following linkage overall and by patient characteristic, route to diagnosis and cancer stage at diagnosis. We then applied the following analyses to address each of our objectives:

1. Prevalence and characteristics of primary care use in the year before diagnosis: we report descriptive analyses of the proportions of patients that visited a GP for any reason 6 and 12 months before diagnosis and those with any 'relevant' symptoms and specific symptoms recorded up to 12 months before diagnosis by diagnostic route.

2. Variations in rates of health-care use between routes to diagnosis: we examined rates of GP consultations; scheduled and unscheduled hospital admissions and A&E visits before diagnosis. A period of 21 months was selected because it was the earliest point for which service use data were complete. We conducted Poisson regression analyses to test the significance between rates, adjusted for patient and area characteristics in two-time periods: 13–21 months before diagnosis (to capture patterns of use likely to be unconnected with cancer); 2–12 months prior to diagnosis (to capture care plausibly within a patient's diagnostic interval cancer but to exclude the diagnostic episode itself). A sensitivity analysis was performed on a subset where records on health-care use were available up to 36 months (supplementary data).

3. Patient and health service factors associated with emergency presentations: we developed logistic regression models, using robust standard errors to account for clustering by practice, to examine patient, area and consultation characteristics associated with emergency routes to diagnosis.

RESULTS

Data set and cohort characteristics. Of the colorectal cancer cases initially identified in the registry data set, 943 (82%) remained after

Table 1. Cohort characteristics		
Characteristic	**N**	**(%)**
Gender		
Female	446	47.30
Male	497	52.70
Age group		
20–59 years	147	15.59
60–69 years	189	20.04
70–79 years	258	27.36
80+ years	349	37.01
Borough		
1	157	16.65
2	340	36.06
3	253	26.83
4	193	20.47
Deprivation quintile		
20% most deprived	170	18.03
20–40%	173	18.35
40–60%	170	18.03
60–80%	173	18.35
20% least deprived	172	18.24
Missing	85	9.01
Diagnostic route		
Emergency	228	24.18
GP urgent/2WW	248	26.30
GP routine/unknown	288	30.54
Consultant/other/unknown	179	18.98
Stage at diagnosis (combined)		
'Early'	264	28.00
'Late'	375	39.77
Missing	304	32.24
GP visits in the 12 months before diagnosis		
Median (25–75 percentiles)	5 (3–8)	
Total		943
Abbreviations: 2WW = 2-week wait; GP = general practitioner.		

matching with local inpatient records. The age profile of cases was similar in the (matched) cohort and the (unmatched) excluded registry cases. There was no difference in the proportion of late-stage cancers but the unmatched group had double the proportion of unstaged cancers (66 *vs* 32%).

The median age at diagnosis in the study sample was 75 years (IQR65-85), which is consistent with national data on the incidence of colorectal cancer by age (Office for National Statistics, 2013). Broadly consistent with the profile of cancers nationally (e.g., Morris *et al*, 2012), 41% of those with staging information had early stage disease but information on staging was missing from 32% of cases in our data set, which is higher than nationally. In our sample, 24% of patients were diagnosed following an emergency presentation, consistent with Elliss-Brookes *et al* (2012), which reported 26% of colorectal cancer tumours were diagnosed following emergency presentations nationally. In just under 19% of cancers, diagnosis did not occur through GP referrals or emergency presentation (Table 1). This group is likely to include patients diagnosed via the Bowel Cancer

Table 2. Prevalence of primary care use and relevant symptoms coded before diagnosis by diagnostic route

| | | Route to diagnosis | | | | | | |
| | Emergency | | GP | | Consultant/ other/unknown | | Total | | |
	n	(%)	n	(%)	n	(%)	N	%	P
Patients with consultations									
In the 6 months before diagnosis	197	84.40	489	91.23	142	79.33	825	87.49	0.008
In the 12 months before diagnosis	203	89.04	508	94.78	156	87.15	867	91.94	<0.001
With any 'relevant' symptom coded within 12 months before diagnosis[a]	165	81.28	481	94.69	121	77.56	768	88.58	<0.001
With specific symptoms coded within 12 months before diagnosis									
Reported anaemia	93	40.79	269	50.19	50	27.93	412		
Anaemia tested but values in normal range or not reported	32	14.04	114	21.27	44	24.58	190		
Constipation	55	24.12	91	16.98	20	11.17	166		
Abdominal pain, swelling, investigation	50	21.93	87	16.23	19	10.61	156		
Rectal bleeding	6	2.63	96	17.91	12	6.70	114		
Diarrhoea	14	6.14	48	8.96	10	5.59	72		
Other bowel (e.g., rectal mass, flatulence, altered bowel)	5	2.19	58	10.82	4	2.23	67		
Weight loss or fatigue	6	2.63	15	2.80	5	2.79	26		
No symptom reported but record of colorectal diagnostic investigation referral	5	2.19	14	2.61	8	4.47	27		
N	228		536		179		943		

Abbreviation: GP = general practitioner.

[a]The percentage is calculated using all those that had a consultation for any reason before diagnosis as the denominator population.

Screening Programme and those under long-term consultant care for bowel disorders.

Prevalence of primary care use and symptoms discussed before diagnosis. Overall, most patients saw their GP within 6 months prior to diagnosis (87% Table 2). The proportions varied significantly between routes in both time periods, with GP referrals more likely than other groups to have seen their GP. Still, 84% of those diagnosed during an emergency presentation saw their GP in the 6 months before diagnosis. The smallest proportion (79%) was in the 'Consultant/other', consistent with the likelihood that this group included cancers detected through screening, and therefore their diagnostic route did not pass through primary care.

Of the patients that saw their GP in the year before diagnosis, most of them (89%) had at least one 'relevant' symptom coded in their notes. The prevalence varied significantly ($P < 0.001$) by diagnostic route with the lowest prevalence of relevant symptoms overall in 'Consultant/other/unknown' (78%) and emergency routes (81%) and highest in the GP-referred group (95%). With respect to specific symptoms, constipation and abdominal pain codes were more prevalent in emergency than GP routes (24 vs 17% and 22 vs 16%, respectively) whereas rectal bleeding was much less common in emergency than GP routes (3 vs 18%). All reported symptoms were less common in the Consultant/other route than other routes.

Rates of health-care use by route to diagnosis. Figure 1 shows the patterns of GP visits (a), scheduled admissions (b), A&E use (c), unscheduled admissions (d) and outpatient attendances (e), up to 21 months before diagnosis for each diagnostic route. Within a year before diagnosis, patterns of GP and urgent care use diverged most widely between groups; those patients diagnosed following a GP referral appeared to see their GP earlier and more frequently than those diagnosed through emergency routes. Rates of A&E visits and emergency admissions rose in the months immediately before diagnosis for all patients but rose most sharply in patients diagnosed through emergency routes. As shown in Table 3, differences in rates of GP visits and urgent care use in the 2–12 months before diagnosis by route were statistically significant, both before and after adjusting for other factors.

As shown in Table 4, patients' service use patterns (i.e., over a time period unlikely to be connected to the route to diagnosis of cancer) also varied significantly by their different diagnostic routes after adjusting for other factors. Patients diagnosed as emergencies had fewer GP visits per month, their rate of urgent care use was over 50% higher (1.56 (95% CI: 1.12; 2.17)) than those diagnosed through GP referrals. In a subset of the cohort where data on service use were available up to 36 months, those diagnosed through emergency routes still had higher emergency care use 12–36 months before their diagnosis, but after 24 months, their primary care use became more similar to other routes (supplementary data).

Factors associated with diagnosis following an emergency presentation. Table 5 shows factors associated with diagnosis following an emergency presentation. Emergency presentations were significantly associated with prior consultation history before and after adjusting for other covariates. The odds of an emergency presentation were lower when patients had seen their GP for any reason. Odds of emergency presentation varied by reported symptoms: they were lower for rectal bleeding (AOR: 0.18 (95% CI 0.07; 0.43)) but higher for abdominal pain/ swelling (AOR: 2.32 (95% CI 1.46; 3.69)) and constipation (AOR: 1.85 (95% CI 1.20; 2.86)).

Emergency diagnoses were also associated with patient, clinical and area factors; they were most common in the oldest group of patients (aged 80 years and over) and least common in those aged 60–69 years. They varied significantly between boroughs; compared with Borough 2, the odds were lower in Borough 3. Those diagnosed with late or missing cancer stage were more likely to be diagnosed as emergencies than those diagnosed at an early stage though this finding was only weakly significant once adjusted for other factors.
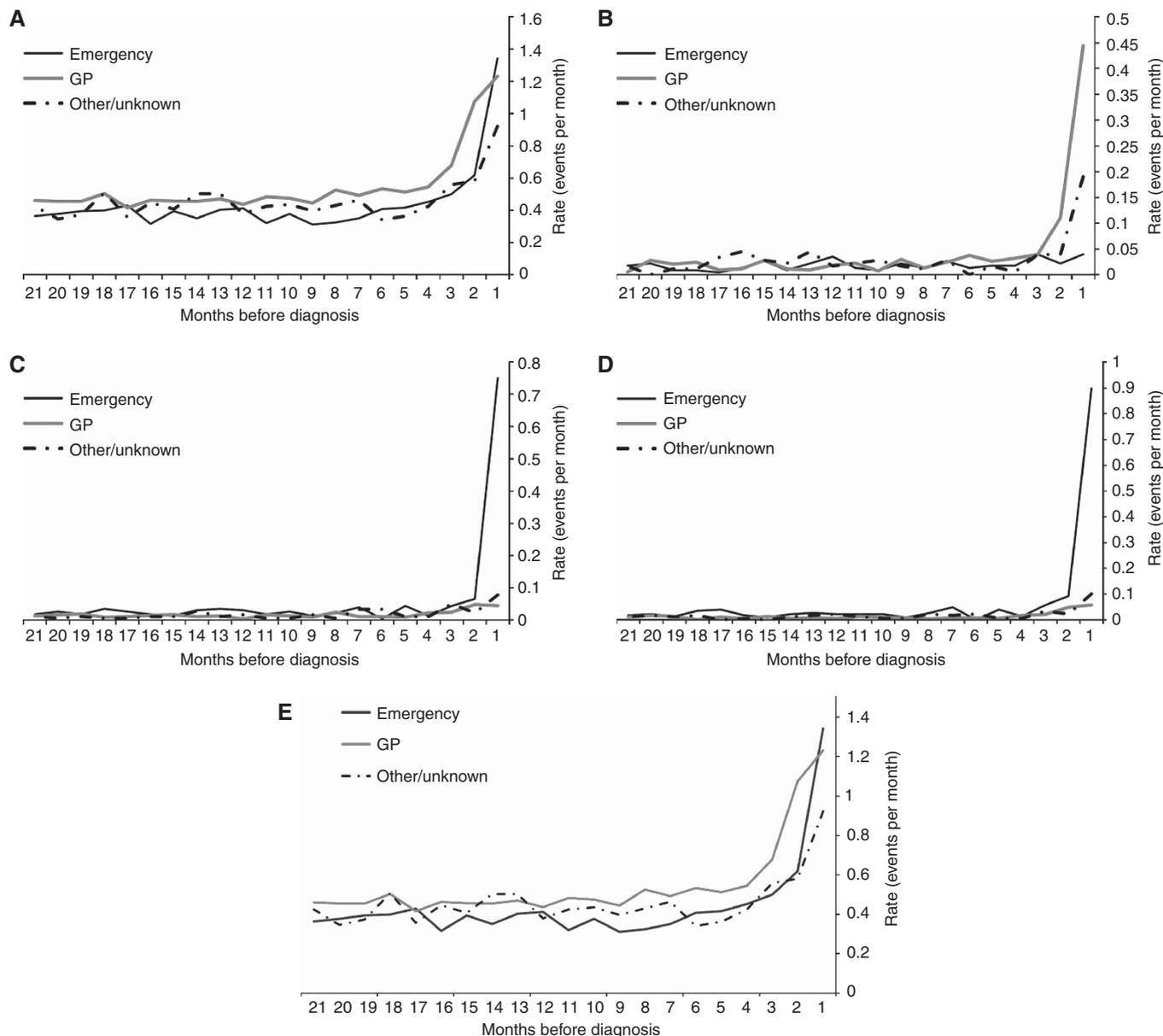
Figure 1. Rates of health service use in the 21 months before diagnosis by route to diagnosis: (**A**) GP visits; (**B**) Scheduled admissions (**C**) A&E visits (**D**) Unscheduled admissions (**E**) Outpatient attendances.

DISCUSSION

Main findings. This study extends examination of the colorectal cancer pathway to encompass patients' first presentation in primary care through to diagnosis in four boroughs of London. Most patients, regardless of the route to diagnosis consulted their GP 6 months before diagnosis but there were distinct patterns of health service use by route. Compared with GP-referred cases, patients diagnosed through emergency routes used urgent care services more frequently and consulted their GP at a later stage. Of those that consulted their GP in the year before their diagnosis, patients diagnosed as emergencies were less likely to have symptoms fulfilling urgent referral criteria coded in their notes.

Strengths and limitations. This study extends previous research in this area by using linked primary, secondary care and cancer registry data, examining patients' health service use in the 21 months before their cancer diagnosis, and focusing on a cancer

where there is the potential to improve survival by reducing emergency presentations. Tsang *et al's* study (2013) used linked primary and secondary care records to examine factors associated with emergency presentations but included health service use only 1 month before any cancer diagnosis. In both Elliss-Brookes *et al* (2012) and Tsang *et al* (2013), all cancer sites were analysed in the same manner. The reasons for emergency presentations vary by cancer site and therefore, these studies can have only limited potential to inform clinical practice and management for individual cancer types. Also, in assigning routes to diagnosis Elliss-Brookes *et al* (2012) needed to make strong assumptions that all hospital activity in the 6 months before cancer diagnosis was related to cancer. They acknowledge 'further site-specific research is required to understand the complex nature of what causes patients to follow their route to diagnosis for each tumour'. In contrast, our focus on one cancer site enabled us to identify and exclude hospital activity with no relevance to a colorectal cancer diagnosis, and therefore provide more refined diagnostic routes.

Table 3. Service use 2–12 months before diagnosis (n = 943)

	Primary care visits: incidence rate ratios						Urgent care (A&E and emergency admissions): incidence rate ratios					
	Unadjusted			Adjusted			Unadjusted			Adjusted		
	Ratio	lower CI	upper CI	Ratio	lower CI	upper CI	Ratio	lower CI	upper CI	Ratio	lower CI	upper CI
Diagnostic route												
GP		1			1			1			1	
Emergency	0.72***	0.68	0.78	0.71***	0.66	0.77	1.88***	1.48	2.40	1.59***	1.23	2.05
Consultant/other/unknown	0.77***	0.72	0.83	0.78***	0.73	0.84	1.03***	0.74	1.41	1.16	0.84	1.60
Gender												
Male		1			1			1			1	
Female	1.07**	1.02	1.13	1.04	0.99	1.10	1.47**	1.17	1.84	1.41*	1.13	1.77
Age group												
20–59 years	0.73***	0.66	0.82	0.77***	0.69	0.86	1.17	0.71	1.92	1.03	0.62	1.69
60–69 years		1			1			1			1	
70–79 years	1.30***	1.19	1.41	1.36	1.25	1.48	1.73**	1.15	2.60	1.69*	1.12	2.54
80 + years	1.35***	1.25	1.46	1.41	1.30	1.53	2.82***	1.94	4.10	2.56***	1.75	3.74
Borough												
1	0.85***	0.78	0.92	0.84**	0.76	0.93	1.12	0.82	1.52	1.16	0.79	1.7
2		1			1			1			1	
3	1.01	0.94	1.08	1.00	0.94	1.07	0.63**	0.45	0.86	0.67*	0.48	0.94
4	0.68***	0.63	0.74	0.68***	0.62	0.75	1.18	0.89	1.57	1.41*	1.01	1.97
Deprivation quintile												
20% most deprived		1			1			1			1	
20–40%	1.14	1.04	1.25	1.06	0.96	1.16	0.88	0.62	1.25	0.89	0.62	1.28
40–60%	1.32	1.20	1.44	1.10	0.99	1.22	0.71	0.49	1.03	0.90	0.59	1.36
60–80%	1.15	1.05	1.26	0.93	0.83	1.03	0.91	0.64	1.29	1.28	0.85	1.93
20% least deprived	1.17	1.07	1.29	0.92	0.82	1.03	0.82	0.57	1.18	1.02	0.66	1.60
missing	0.95	0.84	1.07	0.82**	0.71	0.94	0.92	0.60	1.42	0.82	0.51	1.33
Stage at diagnosis (combined)												
Early		1			1			1			1	
Late/missing	0.94*	0.88	1.00	1.00	1.00	1.00	1.23	0.95	1.60	1.29	0.99	1.68

Abbreviations: CI = confidence interval; GP = Referred by a general practitioner (either urgent/2 week wait or routine/unknown). *P<0.05. **P<0.01. ***P<0.001. All factors significantly associated with the outcome (P = 0.05) were entered into the adjusted models.

This study also illustrates how analysis of linked data sets can inform a local understanding of the quality of processes and outcomes of care for patients with colorectal cancer. Traditional approaches to develop this understanding rely on audits, which are vulnerable to bias due to low and uneven response rates; if those practices least engaged in monitoring or improving the quality of care are least likely to respond, they may underestimate the scale of practice variation. These methods of using routine data could be used to focus which practices and processes require in-depth audit.

The study is subject to several limitations. It was conducted in just one area of England. Although these four boroughs comprise ~1 million people with significant population diversity, they are in general more deprived than average and the experience of cancer patients in London differs from other areas of the country (Saunders et al, 2014). Therefore, patterns of service use and consultation may differ in other populations. However, the patterns of increasing general practice use in this cohort are markedly similar to the rates of general practice use before

diagnosis reported by Hamilton et al (2005) for a cohort in Exeter. In addition, our cohort characteristics were similar to those reported in national studies (e.g., Elliss-Brookes et al, 2012), suggesting it may be appropriate to generalise at least some of our findings beyond this geography.

As in most studies using routine data, incomplete data limited the extent to which we could examine certain key characteristics and the confidence with which we could draw conclusions from our analyses. The inclusion of screen-detected cases and the exclusion of very advanced stage cases with no hospital activity may have biased our results. We were not able to distinguish screen-detected cases from the other diagnostic routes. Using data on screening eligibility, uptake and sensitivity, we estimate ~5% of our sample would have been screen-detected. We assumed these cases would fall predominantly in the 'Consultant/Other' group, given the pathway from a positive screen result to diagnosis does not typically involve a GP referral. The higher proportion of 60–70-year olds (i.e., eligible for screening) in this group and the

Table 4. Health service use 13–21 months before diagnosis (n = 943)

| | Primary care visits: incidence rate ratios | | | | | | Urgent care (A&E and emergency admissions): incidence rate ratios | | | | | |
| | Unadjusted | | | Adjusted | | | Unadjusted | | | Adjusted | | |
	Ratio	lower CI	upper CI	Ratio	lower CI	upper CI	Ratio	lower CI	upper CI	Ratio	lower CI	upper CI
Diagnostic route												
GP	1			1			1			1		
Emergency	0.83***	0.76	0.90	0.85***	0.78	0.93	1.84*	1.35	2.53	1.56**	1.12	2.17
Consultant/other/unknown	0.94	0.86	1.02	0.95	0.88	1.04	0.95	0.62	1.46	1.09	0.71	1.68
Gender												
Male	1			1			1			1		
Female	1.14***	1.07	1.21	1.12**	1.05	1.20	1.43**	1.07	1.91	1.31	0.97	1.76
Age group												
20–59 years	0.59***	0.51	0.67	0.60***	0.53	0.69	1.45	0.74	2.84	1.28	0.65	2.53
60–69 years	1			1			1			1		
70–79 years	1.13*	1.03	1.25	1.20***	1.09	1.32	2.47**	1.42	4.32	2.31**	1.32	4.05
80 + years	1.27***	1.16	1.39	1.30***	1.19	1.43	3.28	1.93	5.57	2.80***	1.63	4.80
Borough												
1	0.85**	0.77	0.94	0.77***	0.69	0.87	1.41	0.97	2.06	0.91	0.56	1.45
2	1			1			1			1		
3	1.24***	1.15	1.34	1.22***	1.12	1.32	0.51**	0.32	0.79	0.56*	0.35	0.89
4	0.76***	0.69	0.84	0.73***	0.65	0.81	1.15	0.79	1.67	0.91	0.59	1.42
Deprivation quintile												
20% most deprived	1			1			1			1		
20–40%	1.13*	1.02	1.26	0.97	0.87	1.09	0.69*	0.46	1.03	0.63*	0.42	0.96
40–60%	1.26***	1.13	1.40	0.96	0.85	1.08	0.30***	0.17	0.51	0.33***	0.18	0.59
60–80%	1.21**	1.08	1.34	0.88*	0.78	1.00	0.40**	0.24	0.64	0.46**	0.26	0.80
20% least deprived	1.00	0.89	1.12	0.73***	0.64	0.83	0.47**	0.30	0.74	0.42**	0.24	0.74
missing	0.92	0.79	1.06	0.71***	0.61	0.83	0.74	0.45	1.22	0.63	0.36	1.11
Stage at diagnosis (combined)												
Early	1			1			1			1		
Late/missing	0.89**	0.83	0.96	0.96	0.90	1.03	0.74*	0.54	1.00	0.74	0.54	1.02

Abbreviations: CI = confidence interval; GP = GP-referred (either urgent/2 week wait or routine/unknown). *P<0.05. **P<0.01. ***P<0.001. All factors significantly associated with the outcome (P=0.05) were entered into the adjusted models.

lower prevalence of any relevant reported symptom compared with the other routes are consistent with this assumption. When the logistic regression model for the odds of an emergency presentation is re-run excluding the 'Consultant/Other' group, the observed associations remain significant, suggesting that low prevalence of symptoms within this group is not responsible for the different patterns of symptom prevalence between emergency routes and other cases.

We excluded 18% of cancer registry records from the study cohort because they did not match records in local hospital inpatient data. Lack of matching could occur where cases had very advanced cancer at diagnosis and died before any inpatient hospital activity could be delivered. These cases, where mortality occurs very shortly after diagnosis, are more likely to be older, emergency presentations and with late stage at diagnosis (Downing et al, 2013). Our examination of matched and unmatched records indicated the age profiles were not significantly different but there were double the proportion of cases with missing cancer stage. Therefore, it is likely our sample under-represents late-stage

emergency presentations and as a result, our analysis may underestimate the strength of associations between late-stage cancer at diagnosis and emergency presentation.

Our list of Read Codes is unlikely to be sufficiently exhaustive to capture every instance where a relevant symptom was discussed with a GP. The prevalence of some of the Read Code-reported symptoms, for example, weight loss, are lower than reported in studies where symptoms were extracted from routine data using medcodes or from GP notes (Cleary et al, 2007; Marshall et al, 2011; Rubin et al, 2011). We did not have access to patient notes to ascertain whether Read Codes were a suitable proxy for the presence or absence of symptoms discussed. However, ascertainment of symptoms, even with detailed notes, may still not be conclusive; patient and GP recall of the first presentation of cancer symptoms do not always align, particularly for vague symptoms (Larsen et al, 2014). We have therefore been cautious in our interpretation of these data, particularly where symptom prevalence was much lower than would be expected.

Table 5. Frequency and odds of an emergency presentation (n = 943)								
	Frequency of emergency presentation		Odds ratio (unadjusted)			Odds ratio (adjusted for all factors associated in unadjusted analysis at P = 0.05)		
Characteristic	n	%	Ratio	Lower CI	Upper CI	Ratio	Lower CI	Upper CI	N
Total no. GP visits (12 m before diagnosis)			0.94**	0.9	0.99	0.94**	0.91	0.98	
'Relevant' symptom coded?									
Reported anaemia: yes	93	22.57	0.86	0.63	1.16		n/a		412
No				1.00					
Rectal: yes	6	5.26	0.15	0.07	0.35	0.18***	0.07	0.43	114
No				1.00			1.00		
Abdominal: Yes	50	32.05	1.61	1.11	2.35	2.32***	1.46	3.69	156
No				1.00			1.00		
Constipation: yes	55	33.13	1.73	1.20	2.49	1.85**	1.20	2.86	166
No				1.00			1.00		
Gender									
Male	112	22.54		1			n/a		497
Female	116	26.01	1.21	0.89	1.64				446
Age group									
20–59 years	36	24.49	2.46**	1.38	4.41	2.28*	1.18	4.42	147
60–69 years	22	11.64		1			1.00		189
70–79 years	49	18.99	1.78*	1.03	3.06	1.78*	0.99	3.18	258
80+ years	121	34.67	4.03***	2.45	6.62	4.70***	2.62	8.42	349
Borough									
1	51	32.48	1.32	0.87	1.99	1.71	0.89	2.70	157
2	91	26.76		1			1.00		340
3	43	17.00	0.56**	0.37	0.84	0.52**	0.34	0.88	253
4	43	22.28	0.78	0.52	1.19	0.81	0.44	1.27	193
Deprivation quintile									
20% most deprived	41	24.12		1			1.00		170
20–40%	43	24.86	1.04	0.63	1.72	1.3	0.72	2.34	173
40–60%	25	14.71	0.54*	0.32	0.92	0.83	0.43	1.63	170
60–80%	38	21.97	0.89	0.56	1.40	1.55	0.82	2.94	173
20% least deprived	41	23.84	0.98	0.63	1.54	1.28	0.66	2.48	172
Missing	40	47.06	2.8***	1.58	4.96	3.51***	1.72	7.19	85
Stage at diagnosis (combined)									
'Early'	50	18.94		1			1.00		264
'Late'/missing	178	26.22	1.52*	1.07	2.17	1.52*	1.01	2.28	679

Abbreviations: CI = confidence interval; GP = Referred by a general practitioner (either urgent/2 week wait or routine/unknown). *P < 0.05. **P < 0.01. ***P < 0.001. All factors significantly associated with the outcome (P = 0.05) were entered into the adjusted model.

We were not able to examine the prevalence of comorbidities across our cohort, which could explain the higher use of emergency services in the 13–21 months before diagnosis in those diagnosed through emergency routes. However, our analysis was adjusted for age, which would be highly correlated with comorbidity. Also, if patients diagnosed through emergency routes also have higher prevalence of comorbidities, higher use of primary care would be expected too. However, primary care use was lower in the 13–21 months before diagnosis than among GP-referred patients.

Comparisons with other studies/possible explanations for our findings. McPhail et al (2013) among others have discussed the theory that poorer survival among those emergency presentation may be because of specific clinical characteristics of their cancers. This theory is supported by empirical evidence from McArdle and Hole (2004), who found more tumours diagnosed through emergency routes were located in the left side of the colon than cancer diagnosed through other routes. Cleary et al in 2007 reported a comparatively high prevalence of abdominal pain, loss of weight and diarrhoea in patients diagnosed following emergency presentations. Cleary et al (2007) study was conducted before 2005 and the introduction of guidance from National Institute for Health and Care Excellence (NICE), which recommended urgent referral for change in bowel habit, diarrhoea, rectal bleeding and

anaemia (National Institute for Health and Care Excellence, 2005). In our cohort diagnosed after the introduction of NICE guidance, the prevalence of abdominal pain was still comparatively high in this group as was constipation, both symptoms that do not align with NICE criteria qualifying for urgent referral. In contrast, there was lower prevalence of symptoms that would meet criteria for referral (rectal bleeding, altered bowel). Our data are also consistent with recent evidence that diagnostic intervals are shorter when NICE-qualifying symptoms are present (Neal et al, 2014). Alternatively, it is possible that patients diagnosed as emergencies were as likely to have these symptoms but either they were less likely to report them or GPs were less likely to record them as Read Codes compared with patients diagnosed by other routes. Although there is plenty of evidence that patients do not report all relevant symptoms and that Read Codes do not capture everything discussed in a consultation, there is no obvious reason why patients or GPs' behaviour would vary between diagnostic routes, making this explanation less likely.

Our data provide some suggestive evidence that patients diagnosed through different routes have different patterns of service use, which is unconnected to cancer. A review of the evidence base to inform urgent care strategy concluded that in urban areas with poor primary care access such as the location of our study, 'many may use an A&E department as their first point of urgent and emergency care' (NHS England, 2013, p11). In our study, patients diagnosed through emergency routes presented to primary care later and used urgent care services more often than patients diagnosed through other routes. However, there were also strong borough-wide differences; emergency presentations were consistently higher in Borough 1 than Boroughs 2 and 3, whereas primary care use was lower. Therefore, in this particular borough, difficulties accessing primary care may drive some patients to present in A&E more readily than other boroughs. Differences in the social profile of boroughs are unlikely to explain these differences, given that cancers diagnosed through emergency routes in our sample were not more common in areas with high socioeconomic deprivation, even though A&E use was higher in cancer patients living in more deprived areas. Differences in individual practice or GP performance are also unlikely explanations for the borough-wide differences, given that diagnosis of colorectal cancers are rare events for GPs; in the 3 years of data in this study, the median number of cases per practice was just four with 18 practices having no cases at all (data not shown).

Consistent with other literature, our study found socio-demographic differences in the profile of patients diagnosed through emergency routes (Bottle et al, 2012). In particular, patients diagnosed through emergency routes were more likely to be 80 years and over. In contrast to previous studies, we found the lowest risk of emergency presentations was in those aged 60–70 years. Our data were collected after the National Bowel Screening Programme was operational, (Morris et al, 2012) so this finding is consistent with the explanation that a higher proportion of cancers in this age group are now detected by screening.

IMPLICATIONS AND AREAS FOR FURTHER RESEARCH

This study confirms that in this population, most patients sought primary care in the months before their diagnosis. Since the publication of guidelines to raise awareness of common symptoms of cancer, diagnostic intervals for colorectal cancer have fallen overall (Neal et al, 2014). However, our study provides suggestive evidence that symptoms are either less likely to be present or recognised in patients that are ultimately diagnosed as emergencies. Sensitising GPs to atypical presentations of cancer symptoms, therefore, may contribute to reduce the prevalence of

emergency presentations. In addition, this finding has methodological implications for calculating the diagnostic interval (defined as first presentation to diagnosis, Weller et al, 2012). The first recording of a relevant symptom in primary care will capture only some of the cancers diagnosed as emergencies even though their cancer journey may well have started in primary care. It then raises the question of how to assign a diagnostic interval to these patients.

Although most patients seek primary care before diagnosis, those presenting as emergencies show different patterns of primary and secondary care use dating from well before their diagnosis, suggesting there may be some scope to reduce emergency presentations through general strategies aimed at discouraging patients' A&E use and encouraging primary care use. However, the borough-wide differences identified in this study indicate that local reviews of primary care access and urgent care use are also needed to identify factors responsible.

We provide here an example of how linked data sets could be used to identify factors for investigation and action to improve survival in colorectal cancer in one local area. There is now potential to establish and validate the methods and applications on other cancers and data sets.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

JS had full access to the data used in this study, drafted the paper and had final responsibility for the decision to submit for publication. JS, TG and XAC conducted analyses of the data. MB contributed to analysis and interpretation of the data and made critical revisions on the paper. All authors commented on drafts of the paper and approved the final version. JS undertook the analysis and data collation while employed as a public health registrar by the London Deanery.

REFERENCES

Bottle A, Tsang C, Parsons C, Majeed A, Soljak M, Aylin P (2012) Association between patient and general practice characteristics and unplanned first-time admissions for cancer: observational study. Br J Cancer 107: 1213–1219.

Cleary J, Peters TJ, Sharp D, Hamilton W (2007) Clinical features of colorectal cancer before emergency presentation: a population-based case-control study. Fam Pract 24: 3–6.

Dixon J, Bardsley M (2012) Predictive risk modelling using routine data: underexploited potential to benefit patients. J Health Serv Res Policy 17: 131–132.

Downing A, Aravani A, Macleod U, Oliver S, Finan PJ, Thomas JD, Quirke P, Wilkinson JR, Morris EJ (2013) Early mortality from colorectal cancer in England: a retrospective observational study of the factors associated with death in the first year after diagnosis. *Br J Cancer* **108**: 681–685.

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**: 1220–1226.

Hamilton W (2012) Emergency admissions of cancer as a marker of diagnostic delay. *Br J Cancer* **107**: 1205–1206.

Hamilton W, Green T, Martins T, Elliott K, Rubin G, Macleod U (2013) Evaluation of risk assessment tools for suspected cancer in general practice: a cohort study. *Br J Gen Pract* **63**: e30–e36.

Hamilton W, Round A, Sharp D, Peters TJ (2005) Clinical features of colorectal cancer before diagnosis: a population-based case-control study. *Br J Cancer* **93**: 399–405.

Larsen MB, Hansen RP, Sokolowski I, Vedsted P (2014) Agreement between patient-reported and doctor-reported patient intervals and date of first symptom presentation in cancer diagnosis - a population-based questionnaire study. *Cancer Epidemiol* **38**: 100–105.

Marshall T, Lancashire R, Sharp D, Peters TJ, Cheng KK, Hamilton W (2011) The diagnostic performance of scoring systems to identify symptomatic colorectal cancer compared to current referral guidance. *Gut* **60**: 1242–1248.

McArdle CS, Hole DJ (2004) Emergency presentation of colorectal cancer is associated with poor 5-year survival. *Br J Surg* **91**: 605–609.

McLennan D, Barnes H, Noble M, Davies J, Garratt E, Dibben C (2011) The English Indices of Deprivation 2010. London, United Kingdom: Communities and Local Government. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf (accessed 01/04/2014).

McPhail S, Elliss-Brookes L, Shelton J, Ives A, Greenslade M, Vernon S, Morris EJ, Richards M (2013) Emergency presentation of cancer and short-term mortality. *Br J Cancer* **109**: 2027–2034.

Morris EJ, Whitehouse LE, Farrell T, Nickerson C, Thomas JD, Quirke P, Rutter MD, Rees C, Finan PJ, Wilkinson JR, Patnick J (2012) A retrospective observational study examining the characteristics and outcomes of tumours diagnosed within and without of the English NHS Bowel Cancer Screening Programme. *Br J Cancer* **107**: 757–764.

National Institute for Health and Care Excellence (2005) *Referral guidelines for suspected cancer*. Available at http://guidance.nice.org.uk/CG27 (accessed 012/06/2014).

Neal RD, Din NU, Hamilton W, Ukoumunne OC, Carter B, Stapley S, Rubin G (2014) Comparison of cancer diagnostic intervals before and after implementation of NICE guidelines: analysis of data from the UK General Practice Research Database. *Br J Cancer* **110**: 584–592.

NHS England (2013) *High quality care for all, now and for future generations: Transforming urgent and emergency care services in England. The Evidence Base from the Urgent and Emergency Care Review*. Available at http://www.england.nhs.uk/ourwork/pe/uec-england/ (accessed: 01/04/2014).

Office for National Statistics (2013) Cancer Statistics Registrations, England (Series MB1)—No. 42, 2011. Available at http://www.ons.gov.uk/ons/rel/vsob1/cancer-statistics-registrations–england–series-mb1-/index.html (accessed: 01/04/2014).

Rubin G, McPhail S, Elliott K (2011) *National Audit of Cancer Diagnosis in Primary Care*. Royal College of General Practitioners: London.

Saunders CL, Abel GA, Lyratzopoulos G (2014) What explains worse patient experience in London? Evidence from secondary analysis of the Cancer Patient Experience Survey. *BMJ Open* **4**: e004039.

Sheather J, Brannan S (2013) Patient confidentiality in a time of care.data. *BMJ* **347**: f7042.

Tate AR, Martin AG, Murray-Thomas T, Anderson SR, Cassell JA (2009) Determining the date of diagnosis–is it a simple matter? The impact of different approaches to dating diagnosis on estimates of delayed care for ovarian cancer in UK primary care. *BMC Med Res Methodol* **9**: 42.

Taylor M, Radford G (2012) *Evaluation of the Bowel Cancer Awareness Pilot in the South West and East of England: 31 January to 18 March 2011*. Available at www.gov.uk/government/uploads/system/uploads/attachment_data/file/215106/dh_133125.pdf (accessed 01/04/2014).

Tsang C, Bottle A, Majeed A, Aylin P (2013) Cancer diagnosed by emergency admission in England: an observational study using the general practice research database. *BMC Health Serv Res* **13**: 308.

Weller D, Vedsted P, Rubin G, Walter FM, Emery J, Scott S, Campbell C, Andersen 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.

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