npg

Routes to diagnosis for cancer – determining the patient journey using multiple routine data sets

L Elliss-Brookes[1], S McPhail*,[2], A Ives[3], M Greenslade[3], J Shelton[2], S Hiom[4] and M Richards[5]

[1]Avon, Somerset, and Wiltshire Cancer Services, South Plaza, Marlborough Street, Bristol BS1 3NX, UK; [2]National Cancer Intelligence Network, 18th Floor, Portland House, Bressenden Place, London SW1E 5RS, UK; [3]South West Public Health Observatory, Grosvenor House, 149 Whiteladies Road, Bristol BS8 2RA, UK; [4]Cancer Research UK, Angel Building, 407 St John Street, London EC1V 4AD, UK; [5]National Cancer Action Team, 18th Floor, Portland House, Bressenden Place, London SW1E 5RS, UK

BACKGROUND: Cancer survival in England is lower than the European average, which has been at least partly attributed to later stage at diagnosis in English patients. There are substantial regional and demographic variations in cancer survival across England. The majority of patients are diagnosed following symptomatic or incidental presentation. This study defines a methodology by which the route the patient follows to the point of diagnosis can be categorised to examine demographic, organisational, service and personal reasons for delayed diagnosis.

METHODS: Administrative Hospital Episode Statistics data are linked with Cancer Waiting Times data, data from the cancer screening programmes and cancer registration data. Using these data sets, every case of cancer registered in England, which was diagnosed in 2006–2008, is categorised into one of eight 'Routes to Diagnosis'.

RESULTS: Different cancer types show substantial differences between the proportion of cases that present by each route, in reasonable agreement with previous clinical studies. Patients presenting via Emergency routes have substantially lower 1-year relative survival.

CONCLUSION: Linked cancer registration and administrative data can be used to robustly categorise the route to a cancer diagnosis for all patients. These categories can be used to enhance understanding of and explore possible reasons for delayed diagnosis.

British Journal of Cancer (2012) **107,** 1220–1226. doi:10.1038/bjc.2012.408 www.bjcancer.com
Published online 20 September 2012
© 2012 Cancer Research UK

Keywords: diagnosis; pathways; referral; survival; emergency; route

Improving cancer survival is a key challenge identified in 'Improving Outcomes: A Strategy for Cancer' (Department of Health, 2011). Cancer survival estimates in England currently fall below those in many European countries. If cancer survival in England was comparable to the European average, then 5000 or more deaths within 5 years of diagnosis could be avoided (Abdel-Rahman et al, 2009; Richards, 2009). The observed lower survival in the first year after diagnosis in England can largely be interpreted as evidence of later diagnosis compared with Europe (Thomson and Forman, 2009). Studies comparing England, Norway and Sweden have also identified a higher number of excess deaths in England, predominantly within the first year of diagnosis, which mainly occur in older patients (Holmberg et al, 2010; Møller et al, 2010; Morris et al, 2011). Later, diagnosis can be caused by delays in presentation, primary care delay, delays between primary and secondary care, and secondary care delay (Rubin et al, 2011). The National Awareness and Early Diagnosis Initiative announced in the Cancer Reform Strategy (Department of Health, 2007) aims to coordinate and provide support to activities and research that promote the earlier diagnosis of cancer. Identifying and categorising the routes taken by patients to their cancer diagnoses will reveal any survival differences across

different presentation routes and help our understanding of how patients with poor prognosis enter secondary care. This could inform targeted implementation of awareness and early diagnosis initiatives, and enable assessment of their success. Identifying different routes for patients will also enable further specific research to be undertaken on a cancer type by type basis to improve understanding of cancer presentation as well as helping to focus improvements in service delivery for patients with poor prognosis.

Previous studies of routes to diagnosis have mainly focussed on the impact of the Two-Week Wait (TWW) referral system introduced in 2000 (whereby patients being urgently referred for suspected cancer by their GP can expect to be seen by a specialist within 2 weeks). They examined patient cohorts at a single secondary care unit or geographically clustered GP practices (Barrett and Hamilton, 2005, 2008; Blick et al, 2010), or review such studies (Thorne et al, 2009). Overall, previous studies show variation in route to diagnosis by cancer type, but also consistently show a large fraction of cases not following routine, urgent or TWW GP Referral routes.

This study explores the feasibility of using routinely collected data to evaluate how patients resident in England diagnosed with malignant cancers between 2006 and 2008 (739667 tumours) accessed secondary care for cancer diagnosis, and whether these 'Routes to Diagnosis' are associated with differences in survival.

Clinical Studies

MATERIALS AND METHODS

A 'Route to Diagnosis' is defined as the sequence of interactions between the patient and the health care system, which lead to a diagnosis of cancer, based on the setting of diagnosis, the pathway and the referral route into secondary care. In many cases, this route begins with a GP consultation. Currently available data limits the portion of the route, which is observable in national data sets, to that within the screening service and secondary care, although this does include the referral method from primary care.

Many routes involve multiple interactions with different parts of the health care system. A large number of individual routes can be defined by combining the setting of diagnosis, the pathway and the referral route, with 71 distinct combinations identified in this cohort. To be useful for analytical purposes, these must be aggregated into a manageable number of broader categories.

Upon examination, two categories were identified, which represent qualitatively different routes (Screen-Detected and Death Certificate Only (DCO)). Three routes reflect the urgency of referral (Emergency, TWW Referral and other GP Referral). Two further routes represent cases for which the route apparently started in secondary care (Inpatient Electives and Other Outpatients) and, finally, one reflects cases with no useful information available on the route to diagnosis (Unknowns). These eight groups are detailed in Table 1.

Cancer registration records for all newly diagnosed malignant tumours excluding non-melanoma skin cancer (ICD-10 C00-C97 excluding C44) diagnosed between 2006 and 2008 in residents of England were extracted from the National Cancer Data Repository (NCDR; National Cancer Intelligence Network, 2010). These cancer registration records were linked at patient level to the administrative Inpatient and Outpatient HES data sets from 2003–2004 to 2008–2009 (NHS Information Centre for Health and Social Care, 2011); the National Cancer Waiting Times (CWT) Monitoring data set from November 2005 to January 2009 (Information Standards Board, 2002); National Breast Screening Programme data from 2005 to 2008 (Association of Breast Surgery, 2011); and National Bowel Screening Programme data from 2006 to 2008 (NHS Bowel Cancer Screening Programme, 2011). These data sets were linked using the unique NHS number that is assigned to each patient in England and which is present in nearly every patient level record in each of the data sets (completeness greater than 98.5% in all data sets, except outpatient HES data which could not be directly assessed). The gynaecological screening status recorded within the NCDR provided screening identification for cervical tumours.

The NCDR data set was deduplicated using European Network of Cancer Registries (ENCR) criteria (Parkin et al, 1994), removing 7.0% of cases.

The Routes to Diagnosis algorithm first used HES data to categorise the route for each tumour individually. National Screening Programme and CWT data linked by NHS number to the cancer registration record were then examined with the assignment of route potentially changing to either a 'Screening' or 'TWW' route.

Figure 1 shows the categorisation of each case into a route using HES data. A specific inpatient or outpatient episode was identified in HES as the 'end point' of the route by its proximity to the date of diagnosis (defined by standard registration rules using ENCR criteria (Tyczynski et al, 2003)). The end point was assumed to be the clinical care event that led most immediately to diagnosis. Having defined the end point, the algorithm seeks a start point of the route. The start point is determined by working backwards from the end point as shown in Figures 1–3 and varies both in the care setting and in the length of time before diagnosis. The characteristics of this start point lead to categorisation of route.
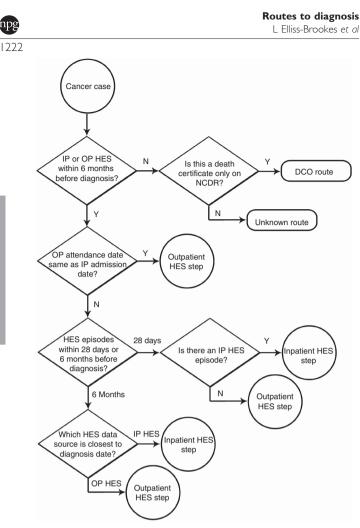
Where both inpatient and outpatient activity occurred on the date of diagnosis, the inpatient episode was defined as the end point of the route. Otherwise, if there was an episode within 28 days before the date of diagnosis, then this was assigned as the end point of the route, with inpatient episodes taking precedence over outpatient episodes, and the most recent episode taking precedence if there were multiple episodes. If there was no HES activity within 28 days of diagnosis, then the most recent episode within 6 months (inpatient or outpatient) was used as the end point of the route. For cases with no HES activity in the 6 months before date of diagnosis, the route was classified as Unknown, or as DCO for cases recorded on the NCDR as being assigned DCO status by cancer registries.

Figure 2 shows the steps taken to seek a start point to the route when the end point was an inpatient admission. The method of admission of the end point was examined to determine the preceding step in the route. Where the method of admission was emergency in nature, this episode was defined as the start point of the route (as well as the end point) and the route was classified as an Emergency Presentation. Where the method of admission was a transfer, the most recent inpatient activity in the 6 months before the admission was examined in an iterative fashion. Where the method of admission indicated a previous outpatient attendance, the most recent outpatient activity in the 6 months before the admission was identified and the source of referral of this

Table 1 The eight routes used to categorise all tumours

Diagnosis route	Description	Priority	Relevant codes
Screen-Detected	Detected via the breast, cervical or bowel screening programmes.	1	Outpatient source of referral: 17
TWW	Urgent GP Referral with a suspicion of cancer.	2 or 4[a]	CWT priority type 1 (to December 2008) CWT priority type 3 (from January 2009)
Emergency Presentation	An emergency route via A&E, emergency GP Referral, emergency transfer, emergency consultant outpatient referral[b], emergency admission or attendance.	3	IP method of admission: 21, 22, 23 and 28 OP source of referral: 1, 4 and 10
GP Referral	Routine and urgent referrals where the patient was not referred under the TWW referral route.	5	Outpatient source of referral: 3 and 12
Inpatient Elective	Where no earlier admission can be found before admission from a waiting list, booked or planned.	5	IP method of admission: 31, 32, 82, 83, 84 and 89
Other Outpatient	An elective route starting with an outpatient appointment: either self-referral, consultant to consultant, other or unknown referral.	5	OP source of referral: 2, 5, 6, 7, 8, 11, 13, 14, 15, 16, 92, 93, 97 and 99
DCO	No data available from Inpatient or Outpatient HES, CWT, Screening and with a death certificate diagnosis flagged by the registry in the NCDR.	5	NA
Unknown	No data available from Inpatient or Outpatient HES, CWT, Screening.	5	NA

Abbreviations: A&E = Accident and Emergency; CWT = Cancer Waiting Times; DCO = Death Certificate Only; GP = general practitioner; HES = Hospital Episode Statistics; IP = Inpatient; NA = not applicable; NCDR = National Cancer Data Repository; OP = Outpatient; TWW = Two-Week Wait. The priority given to each route and relevant codes from each data source are shown. [a]If a TWW record exists, and HES indicates an Emergency route, the TWW takes priority if the emergency admission date is greater than 28 days before the decision to treat date. [b]Only if no previous Outpatient HES data are available for this patient.

Figure 1 Flow diagram for allocating the end point of the route using inpatient and outpatient data.
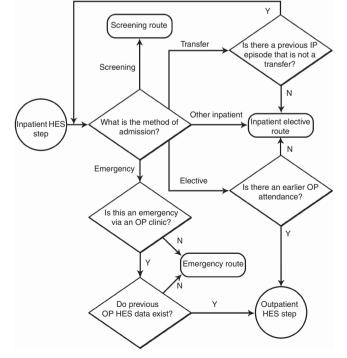

Figure 2 Flow diagram for finding the start point or prior step for an inpatient step in a route.
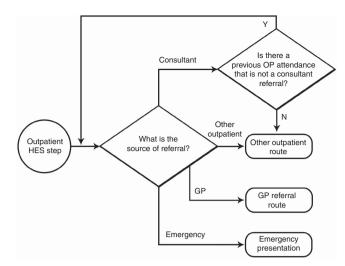

Figure 3 Flow diagram for finding the start point or prior step for an outpatient step in a route.

outpatient attendance examined as described below, except that if there was no outpatient activity within 6 months before the inpatient admission, the route was classified as Inpatient Elective. Otherwise, the route was classified as Unknown or Inpatient Elective according to the codes listed in Table 1.

Figure 3 shows the examination of the outpatient source of referral when the end point of the route was an outpatient attendance. Where the outpatient attendance was not the first appointment of an outpatient episode, the first appointment in that episode was examined. If the source of referral indicated referral from a previous outpatient episode, then the most recent first outpatient attendance within the previous 6 months was examined. Otherwise, the route was classified as Emergency Presentation, Screen-Detected or Other Outpatient according to the codes in Table 1.

After routes were allocated to each case from the HES data, the screening and CWT data were examined. Where a case could be linked to a CWT urgent referral for suspected cancer, it was categorised as a TWW route, unless the route categorised using the HES data was an Emergency Presentation with an admission date within 28 days before the decision to treat date. Where the case could be linked to a screening event, the route was categorised as Screening. If both were possible, then a Screen-Detected route took priority over a TWW route (route priorities are summarised in Table 1).

A case was linked to a CWT referral where a TWW had a decision to treat date within 62 days before or 31 days after the date of diagnosis. A case was linked to a breast screening event where the breast screening assessment date was within 91 days before or 31 days after the date of diagnosis. For colorectal and cervical screening data, the determination that the case was Screen-Detected had been made by the NHS Bowel Cancer Screening Programme (2011) or the regional cancer registries, respectively, and no matching by date was performed.

The algorithm was written in SQL within a Microsoft Server 2005 database environment. Confidence intervals for proportions were calculated using the Wilson score interval (Newcombe, 1998). Calculation of point estimates and confidence intervals for relative survival was done in the statistical package STATA version 10 using the *strel* programme and age-sex-region-deprivation-year lifetables (Cancer Research UK Cancer Survival Group, 2006). Cases were excluded from the analysis if sex, date of birth or date of diagnosis were missing, or if the patient was aged over 99 years at the time of diagnosis. No further stratification by age or case censorship was performed.

RESULTS

The proportion diagnosed by route of the 739 667 tumours categorised is shown in Table 2. Most cancers were diagnosed through one of Emergency Presentation (24%), TWW (26%) or GP Referral (21%) with the other five routes making up 29%. These proportions vary considerably with cancer type.

The proportion of Emergency Presentations increases with increasing age, whereas TWW and GP Referral routes decrease with age. Unknown routes are highest in the under 50 years age group, whereas DCOs are highest in the 85+ years age group. Screening proportions show the effect of the breast screening age range. The proportion of Emergency Presentations amongst children (age 0–14 years) was 54% for all tumours with low TWW proportions (2% overall, data not shown). The proportions for teenagers and young adults (age 15–24 years) were more reflective of the overall cohort, with 24% presenting as an Emergency Presentation for all tumours and higher rates for some sites (e.g., 57% for colorectal, data not shown).

The proportion of routes changed little over the 3 years 2006–2008 (data not shown). For all cancers combined, the proportion categorised as a TWW route increased from 25 to 27%. The proportion categorised as Emergency Presentation was 24% in 2006 and 2007, and 23% in 2008. The proportion of Screen-Detected colorectal cases increased from 0.1% in 2006 to 5% in 2008, reflecting the staged rollout of the NHS Bowel Cancer Screening Programme (2011).

Comparatively small but statistically significant ($P < 0.001$, z-test for difference in proportions) increases in the proportion of tumours diagnosed via a TWW route were seen between 2006 and 2008 for bladder (5%), oropharynx (10%), larynx (7%), melanoma (6%), prostate (5%) and uterine (4%) cancers.

One-year relative survival estimates are presented by route in Table 3, although survival is not calculated for DCO routes (Parkin and Hakulinen, 1991). Across all cancer types, 1-year relative survival was significantly lower for cases categorised as an Emergency Presentation than for those presenting via other routes.

Unknown routes have a comparable survival to other non-Emergency Presentation routes. Where present, the highest relative survival estimates are for Screen-Detected routes.

Sensitivity analysis

The algorithm seeks an inpatient episode within 28 days before diagnosis, and then, any HES episode within 6 months before diagnosis as the end point of the route. Eighty-five percent of cases could be assigned a route from HES data. Of these cases, 74% have a start point within a month before the date of diagnosis. This increases to 84% that have a start point within 3 months and 98% within 6 months before the date of diagnosis.

Overall, 94% of routes to diagnosis were derived either from non-HES data (39%) or from HES data with an end point within 28 days of the day of diagnosis (56%), only 4% of routes were diagnosed from HES data with an end point between 28 days and 6 months before the date of diagnosis.

The frequency of (inpatient) admission in the month before diagnosis is 19 times higher than that in the equivalent month a year before diagnosis (across the whole patient cohort). For persons aged 85+ years, this ratio is 14 : 1.

Tumours (6.2%) were diagnosed in patients with more than one invasive cancer, excluding non-melanoma skin cancer (ICD-10 C00-C97 excluding C44), between 2006 and 2008. If these multiple tumours were excluded, then the overall proportion of Emergency routes increased by less than 0.1% and the overall proportion of Unknown routes increased by 0.2%, other route proportions changed by less than 0.5%. The maximum change in all combinations of route and cancer type on including multiple tumours was 1.7% with a mean absolute change of 0.2%.

If Emergency Presentation routes are given the highest priority, followed by TWWs, Screening and others in that order, then Emergency Presentations rise by 0.6%, TWWs drop by 0.4% and Screening drops by 0.2%. If Screen-Detected routes are given the highest priority, followed by TWWs, Emergency Presentations and

Table 2 Proportion of tumours by route, for selected tumours

	Screen-detected (%)	TWW (%)	GP referral (%)	Other outpatient (%)	Inpatient elective (%)	Emergency presentation (%)	DCO (%)	Unknown (%)	n
All cancers	5	26	21	10	6	24	1	8	739 667
Under 50 years	2	29	24	10	6	15	0	13	81 072
Aged 50–59 years	12	26	21	9	6	15	0	10	102 487
Aged 60–69 years	10	26	22	10	6	18	0	8	181 958
Aged 70–79 years	2	28	23	10	6	25	1	6	207 389
Aged 80–84 years	0	25	20	9	5	34	1	6	87 940
Aged 85+ years	0	20	16	7	4	43	3	7	78 821
Bladder		30	24	13	9	19	1	5	25 639
Central nervous system		1	13	11	7	62	1	6	11 697
Breast	28	43	11	3	1	5	0	9	110 173
Colorectal	2	27	20	9	9	26	1	6	91 416
Kidney and unspecified urinary organs		19	26	17	6	25	1	6	20 594
Lung		24	17	10	4	39	1	5	96 735
Melanoma		41	27	7	3	3	0	18	26 660
Multiple myeloma		11	27	13	6	37	1	6	11 221
Non-Hodgkin lymphoma		18	28	12	6	27	0	9	25 413
Oesophagus		34	16	8	14	22	1	5	19 449
Ovary		23	20	12	5	32	1	7	16 026
Pancreas		11	16	9	6	50	1	6	19 896
Prostate		26	32	11	8	10	0	12	92 922
Stomach		23	17	8	13	33	1	5	18 613
Uterus		37	31	10	5	8	0	8	18 462

Abbreviations: DCO = Death Certificate Only; GP = general practitioner; TWW = Two-Week Wait. Cases diagnosed in persons with an English residential address, 2006–2008. Cervical cancer proportions relate to 2006–2007 data due to incomplete screening data in 2008. All 95% confidence intervals are below ±1%.

Clinical Studies

Table 3 One-year relative survival by route, for selected tumours with 95% confidence intervals

	All routes		Screen-Detected		TWW		GP Referral		Other Outpatient		Inpatient Elective		Emergency Presentation		Unknown	
	Survival	95% CI	Survival	95% CI	Survival	95% CI	Survival	95% CI	Survival	95% CI	Survival	95% CI	Survival	95% CI	Survival	95% CI
Bladder	73%	72–73			83%	82–84	79%	78–80	77%	75–79	83%	81–84	36%	35–37	74%	72–77
Central nervous system	39%	39–40			47%	37–57	54%	52–57	62%	59–65	53%	50–57	30%	29–31	50%	46–54
Breast	97%	96–97	100%	100–100	98%	98–98	96%	96–97	92%	91–93	91%	88–93	54%	52–55	95%	95–96
Colorectal	74%	74–74	98%	97–98	82%	82–83	82%	81–83	80%	79–81	84%	83–85	50%	49–51	73%	72–74
Kidney and unspecified urinary organs	69%	68–70			79%	77–80	80%	79–81	82%	81–84	78%	75–80	38%	37–40	63%	60–66
Lung	29%	28–29			40%	39–41	40%	40–41	44%	43–45	34%	33–36	12%	11–12	24%	23–25
Melanoma	97%	97–97			99%	98–99	98%	97–98	94%	92–95	96%	95–98	62%	58–66	99%	98–99
Multiple myeloma	70%	69–71			82%	80–85	81%	79–82	78%	75–80	79%	76–83	51%	49–53	80%	76–83
Non-Hodgkin lymphoma	75%	75–76			85%	84–87	86%	85–87	81%	79–82	84%	81–86	50%	49–51	86%	84–88
Oesophagus	40%	39–40			42%	41–43	47%	45–48	50%	48–53	49%	47–51	18%	17–20	44%	41–48
Ovary	70%	69–70			84%	82–85	81%	79–82	82%	80–84	81%	78–84	45%	44–47	68%	65–71
Pancreas	17%	16–17			19%	18–21	26%	24–27	33%	31–35	29%	26–32	9%	9–10	16%	14–18
Prostate	96%	95–96			98%	97–98	99%	99–99	96%	96–97	99%	99–99	60%	59–62	98%	98–99
Stomach	41%	40–41			43%	42–45	52%	50–54	55%	52–58	53%	51–55	23%	21–24	44%	41–47
Uterus	91%	90–91			94%	94–95	94%	93–94	90%	89–92	93%	91–95	59%	56–61	89%	87–91

Abbreviation: CI = Confidence Interval. Cases diagnosed in persons with an English residential address, 2006–2008.

others in that order, then Emergency Presentations drop by 1.2%, TWWs rise by 1.2% and Screening is unchanged.

Screening data and CWT data are linked specifically to the cancer record rather than to the patient as for HES data. As such, these data are treated as more robust and, therefore, routes generated from them supersede the route generated from the hospital admissions records, with the exception of Emergency Presentations admitted within 28 days before the decision to treat date. There is an impact of less than 0.2% of cases if Screen-Detected routes were not prioritised above Emergency Presentations.

A TWW record was used to categorise the route if the decision to treat date fell within a 3-month period from 31 days before to 62 days after the diagnosis date. A screening record was used to categorise the route if the screening date fell within a 4-month period from 91 days before to 31 days after the diagnosis date. These periods were chosen to correspond to the typical timescales of these patient pathways, and to take account of cancer registration rules, which preferentially define the date of diagnosis from pathological confirmation. Sensitivity analysis showed that the route breakdowns were not greatly affected by changes of a month in the length of the screening date periods; a reduction of 4% in the proportion of TWW routes was observed if the TWW matching period was reduced to 1 month before to 1 month after the diagnosis date. Prioritising all TWW routes above Emergency Presentations reduces the observed proportion of Emergency Presentations by around 1%.

DISCUSSION

The routes to diagnosis algorithm

A central assumption underlying the algorithm is that it is reasonable to suppose that inpatient and outpatient hospital activity up to 6 months, and in particular in the 28 days before

diagnosis, is linked to the diagnosis of the cancer. This activity may not necessarily be directly caused by the cancer itself as diagnosis can result from other clinical investigations, for example, radiological examination for an unrelated condition.

The higher frequency of activity in the month before diagnosis compared with that a year earlier indicates that the majority of hospital activity in the 28 days before diagnosis is related to the diagnosis of cancer. Making the conservative assumption that a 'background' event has an equal chance of being picked up by the algorithm as events related to the diagnosis, and the further conservative assumption that they will give an incorrect aggregated route allows us to estimate an overall upper bound of 10% on the error rate of the algorithm due to 'background' admissions to hospital. The resultant uncertainty in the proportion of cancers diagnosed via a route with a proportion of 25% of cancers would be approximately 2.5% points overall, and slightly higher for persons aged 85+ years. A small bias toward non-Emergency Presentations might be expected for older persons, because of the fact that the majority of their higher 'background' admission rate are elective admissions. Similarly, small systematic effects in specific patient groups with pre-existing comorbid conditions will exist, with the resulting bias depending on the typical nature of an admission for the comorbid condition.

The algorithm does not attempt to match diagnosed cancers to cancer-specific inpatient or outpatient HES records. The majority of outpatient records do not have diagnostic coding, and even where it does exist in outpatient or inpatient records, it is likely that the episodes of interest (being pre-diagnostic) would not include cancer-specific codes. The algorithm only relies on the existence of attendance and episode records, and the associated administrative fields recording source of referral and method of admission, making the calculated routes insensitive to variation in clinical coding.

The methodology presented has several adjustable parameters. The inclusion of multiple tumours did not substantially affect the

Table 4 Proportion of tumours in selected comparable studies for GP, TWW and Emergency routes

Tumour type	Study	n	Presentation route		
			All GP (%)	TWW (%)	Emergency (%)
Bladder	Blick et al (2010)	100	80	42	15
	Routes to Diagnosis	25 639	54	30	19
Colorectal	Neal et al (2007)	239		21	
	Thorne et al (2009)	1679		33	
	Routes to Diagnosis	91 416	47	27	26
Lung	Neal et al (2007)	409		23	
	Routes to Diagnosis	96 735	41	24	39
Ovarian	Neal et al, 2007	95		24	
	Routes to Diagnosis	16 026	43	23	32
Prostate	Barrett and Hamilton, 2005	217	76		11
	Neal et al (2007)	146		32	
	Routes to Diagnosis	92 922	59	26	10
Upper GI	Thorne et al (2009)	498		34	
	Routes to Diagnosis	66 534	37	21	37

Abbreviations: GI = gastrointestinal tract; GP = general practitioner; TWW = Two-Week Wait.

magnitude of the variation in frequency between routes and they were therefore included in the analysis (Brenner and Hakulinen, 2007; Rosso et al, 2009). Changing the priority of Emergency Presentations with respect to Screening and TWW routes slightly (approximately 1% or less) alters the proportions of cases categorised as each route. The lack of overlap between Emergency Presentation and Screen-Detected routes is reassuring as the majority of Screen-Detected cases are early-stage tumours that are less likely to result in an Emergency Presentation.

Although the methodology allows the assignment of a Route to Diagnosis, it is not intended to identify presenting symptoms of either the cancer or of other illnesses, which may have then led to the cancer diagnosis. Further site-specific research is required to understand the complex nature of what causes patients to follow their Route to Diagnosis for each tumour.

Route frequencies and impact on survival

In every tumour type examined, 1-year relative survival was significantly lower for Emergency Presentations than for any other route. The magnitude of the difference in survival between Emergency and non-Emergency routes is typically in the range of 20–40%. The higher proportion of older people with Emergency Presentations may partly explain this difference in survival. One-year relative survival is lower for the TWW route compared with other non-emergency routes for several cancer sites, including cancers of the central nervous system, liver and lower GI cancers. Given the comparative rapidity of TWW referrals, this could be an example of a waiting time paradox (Crawford et al, 2002). This is consistent with other studies (Tørring et al, 2011) showing that outcomes were worse for the most urgently referred cases.

Cases allocated an Unknown route have a cancer registration (not based solely on a death certificate), but no data can be found in HES in the 6 months before diagnosis, or in CWT or screening sources. The higher proportion of Unknown routes in people under 50 and in the more socio-economically affluent (data not shown) may indicate a higher fraction of private referrals in this group. The effect of age is also seen in the National Audit of Cancer Diagnosis in Primary Care (Rubin et al, 2011), which indicates that private referrals are less common in people over 70. The relatively large proportion (18%) of tumours assigned to the Unknown route for melanoma might be due to tumours being removed in primary care where melanoma was not suspected. The survival of the Unknown routes is comparable to that of other non-emergency routes across all cancer types, suggesting delivery of care via non-emergency settings.

Comparisons with other studies

This study calculates proportions of routes at a population level using nationally defined data sets. When comparing these proportions with previous studies, the nature of the patient cohort should be considered. Patient cohorts from primary care may under-record Screen-Detected cases, incidental diagnosis made while under hospital care and Emergency Presentations that result in death shortly after diagnosis. Patient cohorts from secondary care may under-report clinical diagnoses if case finding relied on histopathological databases. In addition, statistical fluctuations in the observed proportions may occur in studies conducted at single centres because of the comparatively low number of cases diagnosed for each tumour type over the study period.

Comparable results from other studies are presented in Table 4. There is generally a good agreement between the proportion of cases diagnosed via TWW in our study and the majority of those studies in which case finding was done in secondary care (Neal et al, 2007; Thorne et al, 2009; Blick et al, 2010). The proportion of Emergency Presentations seen in this study is also comparable to those seen in studies with case finding via secondary care or cancer registries (Barrett and Hamilton, 2005; Blick et al, 2010).

The total proportion of cases which present via GP Referral is higher in the studies examined compared with this study. This might be explained by the inclusion in this study of separate categories for Screen-Detected, Unknown, Other Outpatient and Inpatient Elective routes. In particular, it is possible that the majority of cases in the Other Outpatient and Inpatient Elective routes were originally initiated by a GP Referral. Further work linking Routes to Diagnosis results to GP data sets is needed to explore this supposition.

Although some of the eight routes are specific to the English health care system; the methodology can be used in other countries where data sets exist, detailing episodes of hospital care. The routes of TWW, GP Referral, Inpatient Elective and Other Outpatient could all be considered to have originated from a GP Referral. Thus, a more general international comparison would be possible using only five distinct routes, with these four forming an overall 'GP-Initiated' route.

In summary, we have demonstrated a methodology for categorising a Route to Diagnosis for all registered tumours, using large routinely available health service data sets. This can be applied in an automated fashion to all patients diagnosed with cancer in England that are recorded by the cancer registries and enables research to be undertaken to understand differences within

Clinical Studies

these groups. The frequencies with which these routes are followed in the diagnosis of cancer are in reasonable agreement with previous clinical studies, and show plausible variation by cancer type and age. The route has a significant association with 1-year relative survival. In particular, the substantially lower relative survival in Emergency compared with non-emergency routes indicates that this distinction is of high clinical significance.

REFERENCES

Abdel-Rahman M, Stockton D, Rachet B, Hakulinen T, Coleman MP (2009) What if cancer survival in Britain were the same as in Europe: how many deaths are avoidable? *Br J Cancer* **101**(Suppl 2): S115–S124

Association of Breast Surgery (2011) *Audit of Screen Detected Cancer.* http://www.cancerscreening.nhs.uk/breastscreen/publications/archived-baso-reports.html (accessed 2 August 2012)

Barrett J, Hamilton W (2005) Pathways to the diagnosis of prostate cancer in a British city. A population-based study. *Scand J Urol Nephrol* **39**: 267–270

Barrett J, Hamilton W (2008) Pathways to the diagnosis of lung cancer in the UK: a cohort study. *BMC Fam Pract* **9**: 31

Blick C, Bailey D, Haldar N, Bdesha A, Kelleher J, Muneer A (2010) The impact of the two-week wait rule on the diagnosis and management of bladder cancer in a single UK institution. *Ann R Coll Surg Engl* **92**: 46–50

Brenner H, Hakulinen T (2007) Patients with previous cancer should not be excluded in international comparative cancer survival studies. *Int J Cancer* **121**: 2274–2278

Cancer Research UK Cancer Survival Group (2006) *strel* computer program, version 1.2.7 and life tables for cancer survival analysis. Downloaded from www.lshtm.ac.uk/ncde/cancersurvival/tools on 8 February 2010; Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, UK

Crawford SC, Davis JA, Siddiqui NA, de Caestecker L, Gillis CR, Hole D (2002) The waiting time paradox: population based retrospective study of treatment delay and survival of women with endometrial cancer in Scotland. *BMJ* **325**: 196

Department of Health (2007) *Cancer Reform Strategy.* Department of Health: London, England

Department of Health (2011) *Improving Outcomes: A Strategy for Cancer.* Department of Health: London, England

Holmberg L, Sandin F, Richards M, Spicer J, Lambe M, Klint A, Peake M, Strand TE, Linklater K, Robinson D, Møller H (2010) National comparisons of lung cancer survival in England, Norway and Sweden 2001-2004: differences occur early in follow-up. *Thorax* **65**(5): 436–441

Information Standards Board (2002) DSC Notice: 22/2002 www.isb.nhs.uk/documents/dscn/dscn2002/222002.pdf (accessed 2 August 2012)

Morris EJ, Sandin F, Lambert PC, Bray F, Klint A, Linklater KM, Robinson D, Pahlman L, Holmberg L, Møller H (2011) A population-based comparison of the survival of patients with colorectal cancer in England, Norway and Sweden between 1996 and 2004. *Gut* **60**(8): 1087–1093

Møller H, Sandin F, Bray F, Klint A, Linklater KM, Purushotham A, Robinson D, Holmberg L (2010) Breast cancer survival in England, Norway and Sweden: a population-based comparison. *Int J Cancer* **127**(11): 2630–2638

Neal R, Allgar V, Nasreen A, Leese B, Heywood P, Proctor G, Evans J (2007) Stage, survival and delays in lung, colorectal, prostate and ovarian cancer: comparison between diagnostic routes. *Br J Gen Pract* **57**: 212–219

Newcombe R (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* **17**: 857–872

National Cancer Intelligence Network (2010) *National Cancer Data Repository.* http://www.ncin.org.uk/collecting_and_using_data/national_cancer_data_repository/default.aspx (accessed 2 August 2012)

NHS Bowel Cancer Screening Programme (2011) http://www.cancerscreening.nhs.uk/bowel/index.html (accessed 2 August 2012)

NHS Information Centre for Health and Social Care (2011) *Hospital Episode Statistics (HES).* http://www.hesonline.nhs.uk (accessed 2 August 2012)

Parkin DM, Hakulinen T (1991) Analysis of survival. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG (eds) *Cancer Registration, Principles and Methods. IARC Scientific Publications No. 95.* IARC Press: Lyon, pp 159–176

Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL (1994) *Comparability and Quality Control in Cancer Registration.* IARC Technical Report No.19. IARC: Lyon

Richards M (2009) The size of the prize for earlier diagnosis of cancer in England. *Br J Cancer* **101**: S125–S129

Rosso S, De Angelis R, Ciccolallo L, Carrani E, Soerjomataram I, Grande E, Zigon G, Brenner H, the EUROCARE Working Group (2009) Multiple tumours in survival estimates. *Eur J Cancer* **45**: 1080–1094

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

Thomson C, Forman D (2009) Cancer survival in England and the influence of early diagnosis: what can we learn from recent EUROCARE results? *Br J Cancer* **101**(Suppl 2): S102–S109

Thorne K, Hutchings H, Elwyn G (2009) The two-week rule for NHS gastrointestinal cancer referrals: a systematic review of diagnostic effectiveness. *Open Colorectal Cancer J* **2**: 27–33

Tørring 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**: 934–940

Tyczynski J, Démaret E, Parkin D (eds) (2003) Standards and guidelines for cancer registration in Europe. *IARC Technical Publication* **40**: 69–73