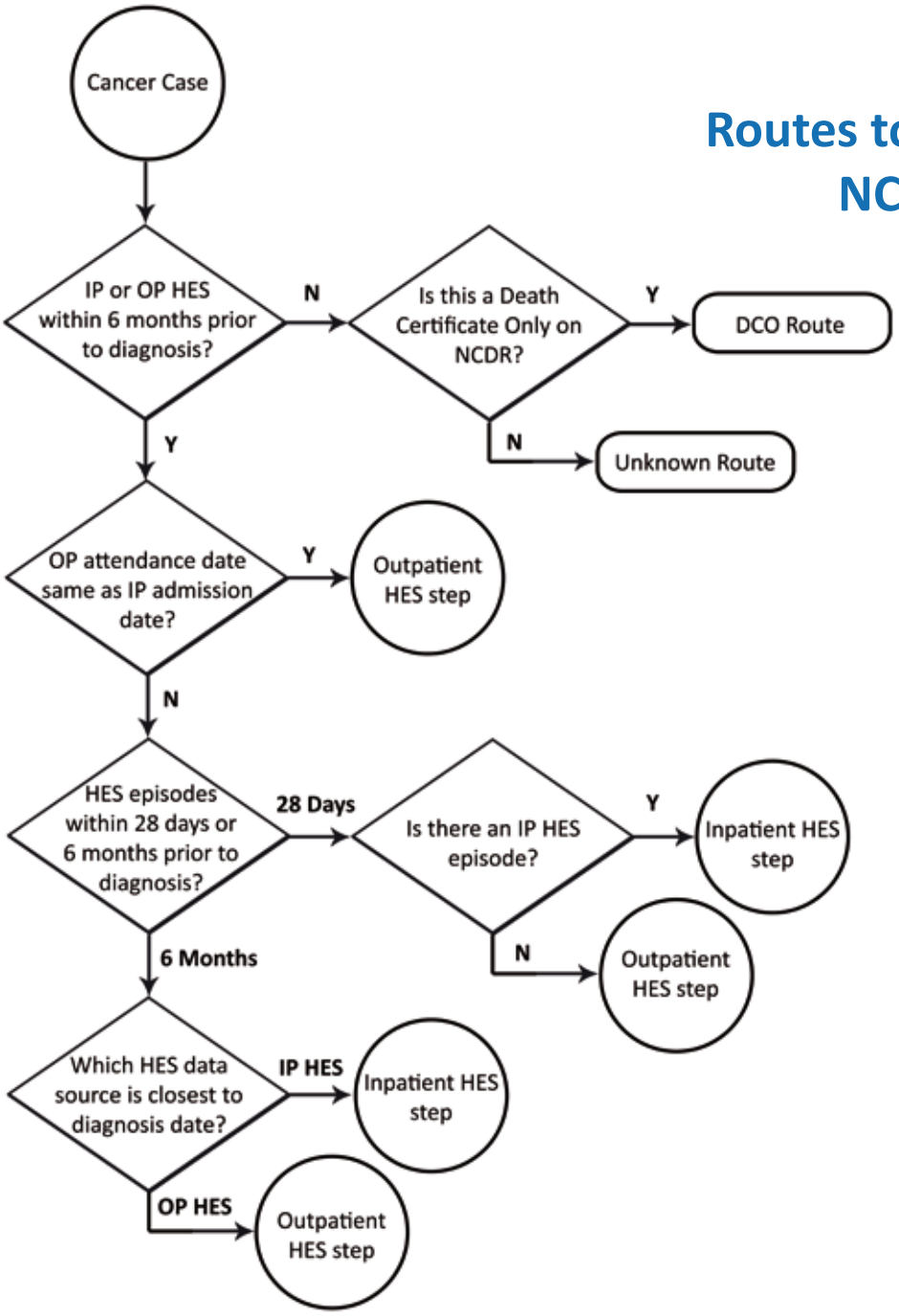
NCIN
national cancer intelligence network
Using information to improve quality & choice



Routes to Diagnosis, 2006-2008
NCIN technical document

Routes to Diagnosis 2006-2008: Technical document

Contents

1.0 Introduction

This document summarises the data sources and methodology used for the second iteration of the 'Routes to Diagnosis' project. Results and analysis are available on the NCIN website[1] and published in the British Journal of Cancer[2].

1.1 Overview of the Routes to Diagnosis project

Project goals

The questions examined in the 'Routes to Diagnosis' project are described below:

- *Is it feasible to use routinely available data sources to define the routes to diagnosis for patients diagnosed with cancer (for example, whether they present through inpatients, outpatients, screening or via an emergency presentation)?*
- *If the first is feasible, can the influence of age, sex, ethnicity, deprivation and geographical area of residence on referral routes and pathways be examined?*
- *Is there an association between routes to diagnosis and survival for cancer patients?*

This document and associated publications[1, 2] demonstrate a positive answer to these three questions.

Technical overview

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

Policy context

The Routes to Diagnosis project supports the National Awareness and Early Diagnosis Initiative (NAEDI) whose aim is to promote earlier diagnosis of cancer and thereby improve survival rates and reduce cancer mortality. Successful implementation of NAEDI will make a major contribution to the Cancer Reform Strategy goal of achieving world class cancer outcomes in this country.

The Routes to Diagnosis project is the first to explore the feasibility of using routine data to evaluate how cancer patients access the health service for diagnosis and whether the routes are associated with survival differences. This in turn could be used to inform strategy in terms of improved patient education regarding signs and symptoms, medical practitioner education, and routes of referral. The outputs will help to inform awareness and early diagnosis initiatives locally and nationally, ideally resulting in more appropriate referrals and earlier diagnosis of cancer as well as eventually improving the cost effectiveness of NHS.

History

The first iteration of the national study was conducted in the summer of 2010 and covered patients diagnosed in 2007[3]. The second iteration refines the algorithm used and widens the period of data analysed to cancers diagnosed in 2006-2008 (inclusive).

2.0 Methods

This section describes the process by which the Routes to Diagnosis algorithm assigns a Route to each cancer recorded in the National Cancer Data Repository.

2.1 Overview of the Routes to Diagnosis algorithm

The algorithm takes as a starting point the date of cancer diagnosis, as defined by the UKACR using European Network of Cancer Registries (ENCR) rules[4]. Routine data immediately prior to this date is examined and a series of rules is used to classify the 'Route to Diagnosis' for each case. The routes are categorised in detail by three variables: the end-point, the pathway group, and the start-point. These detailed routes have been aggregated into eight broader categories to facilitate analysis.

It is important to note that patient records being used to describe the route to diagnosis may not have a cancer code assigned to them, as the episodes and attendances will have taken place before a cancer diagnosis has been coded. It is therefore not possible to be absolutely certain that the episodes and attendances related to the patient prior to diagnosis were directly related to the process of diagnosis of cancer. However the frequency of hospital attendance and admission in the period immediately before diagnosis greatly exceeds the 'background' rate making the assumption that they are related to the cancer diagnosis reasonable[2].

2.2 Data sources

Cancer registration

The National Cancer Data Repository holds cancer registration data for the whole of England. The repository contains over 6.5 million cancer registry records. Further information about the National Cancer Data Repository is available from the NCIN website: www.ncin.org.uk.

All cancer registrations across England between 2004 and 2008 inclusive, with ICD-10 diagnosis codes C00–C97 and D00-D48 (all neoplasms) were obtained from the National Cancer Data Repository.

A subset of this data for tumours with ICD-10 diagnosis codes C00-C97 excluding C44 which were diagnosed in calendar years 2006 to 2008 was used for reporting. Other records were excluded from the reporting dataset based on experience in the first iteration of Routes to Diagnosis:

- The records for patients with ICD-10 codes D05 and D06 (in situ breast and in situ cervical cancers) because it was found that the data for these is of very variable quality with substantial missing information. Other ICD-10 "D codes" were excluded for a similar reason;
- The records for patients with non-melanoma skin cancer, as most of these are diagnosed and treated immediately in outpatients or in primary care and Basal Cell Carcinomas are not subject to the Two Week Wait referral process;
- Records for diagnosis years 2004 and 2005. Breast screening data for 2004 was not available and the NCWT data was incomplete as the central collection of Cancer Waiting Times data was new during that time.

Routes were derived for all tumours fitting the criteria specified, including second and subsequent tumours in the same person (unlike the first iteration of Routes to Diagnosis). Sensitivity analysis showed a small impact on the total if these multiple tumours were excluded: 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%.

Hospital Episode Statistics (HES)

Hospital Episode Statistics (HES) is a data warehouse containing details of all admissions (day case and inpatient) to NHS hospitals in England. It includes details of private patients treated in NHS hospitals, patients who were resident outside England and of care delivered by treatment centres (including those in the independent sector) funded by the NHS. HES also contain details of all NHS outpatient appointments (attendances for patients who are not formally admitted) in England. It contains admitted patient care data from 1989 onwards, with more than 12 million new records added each year, and outpatient attendance data from 2003 onwards, with more than 40 million new records added each year. Further information about HES is available from the HES online website: www.hesonline.org.uk.

Inpatient Hospital Episode Statistics

For the national analysis, Inpatient (IP) Hospital Episode Statistics (HES) for 2003-04 to 2008-09 were used to identify patients with a hospital admission for any cause during this time period.

Outpatient Hospital Episode Statistics

For the national analysis, Outpatient (OP) Hospital Episode Statistics (HES) for 2003-05 to 2008-09 were used.

National Cancer Waiting Times

For the national analysis, National Cancer Waiting Times (NCWT) data for 2004 to 2008 were used. The NCWT system is hosted nationally on NHSNet (Open Exeter) and allows NHS providers to record data derived from patient care activity. These data are used to monitor performance against the NCWT standards specified in the NHS Cancer Plan 2000 and the Cancer Reform Strategy 2007. As a patient moves through the stages of their treatment pathway, data on referrals, treatments and diagnosis are derived from care records locally. NHS providers are mandated by Data Set Change Notice (DCSN) 20/2008 to collect data concerning all patients covered by the NCWT standards, including patients referred with suspected cancer and patients diagnosed with and treated for new and subsequent cancer. Further information about the NCWT system is available from the Department of Health website: www.dh.gov.uk.

Breast Screening

An offload of the National NHS Breast Screening Programme database was provided by the West Midland Cancer Intelligence Unit. This listed screening attendances and subsequent treatment of breast cancer in women with a screening appointment in calendar years 2005 to 2008.

Cervical screening

An extract of the National Cancer Data Repository was sent to each of the English cancer registries. The extract contained the cervical cancer patients diagnosed in 2007 in the relevant registry catchment area. Each registry was asked to complete the screen detected field (indicating whether the tumour was detected via the National Screening Programme), which was then used to update the extract from the National Cancer Data Repository.

Screening data were provided by the cancer registries, based on the records held internally as a result of local data exchanges between the cancer registries and the screening Quality Assurance Reference Centres (QARCs).

Colorectal screening

An offload was provided by the North Yorkshire Cancer Registration and Intelligence Service, based on data received from the NHS Bowel Cancer Screening Programme. This identified cases of colorectal cancer within the National Cancer Data Repository between 2006 and 2008 (inclusive) that were screen detected.

2.3 De-duplication

The NCDR dataset was de-duplicated using European Network of Cancer Registries (ENCR) criteria[5], removing 7.0% of cases.

2.4 Matching algorithms

Matching algorithm for cancer registration data and HES data

Initially, patient records from the National Cancer Data Repository were matched to both inpatient and outpatient HES records using NHS number only for extract purposes. Then a graded system of matching was applied in order to protect against data quality issues. Patient records were matched using four data fields: NHS number, date of birth, postcode of residence, and sex according to any of the following criteria:

1. NHS number, date of birth, postcode and sex (best possible match).
2. NHS number, date of birth and sex for cases not already matched above. This is because there is a chance that people may have changed address between datasets being compiled, especially if the data span a long time frame.
3. NHS number and postcode for cases not already matched above. This attempts to catch anyone whose date of birth details were unknown or incorrectly entered or where there was an incorrect entry for their sex.
4. NHS number only (least reliable match) for cases that could not be matched above.

Matching cancer registration data and National Cancer Waiting Times Data

Records were extracted by Trent Cancer Registry that had a referral priority of Two Week Wait and a valid Decision to Treat Date and matched the NHS Numbers in the NCDR data for years 2004-2008. These were then matched to the records in the National Cancer Data Repository using NHS Number and having a Cancer Diagnosis date between 62 days before and 31 days after the Decision to Treat Date. 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 one month before to one month after the diagnosis date.

2.5 The Routes to Diagnosis Algorithm

The Routes to Diagnosis Algorithm assigns a three part code to each tumour based on the inpatient and outpatient HES data, as described below in sections 2.6-2.9. This three part code is either mapped to one of seven broader route categories or the presence of Screening or Cancer Waiting Times data can take precedence and cause the final route to be a Two Week Wait or Screen Detected Route, as described in section 2.10.

2.6 Assigning the Route end-point

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. The end-point was assumed to be the clinical care event that led most immediately to diagnosis. 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 prior to 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.

The following end-point codes were assigned:

Special cases (SC) – patients with a cancer diagnosis date on the same day as an inpatient admission date and an outpatient attendance date, or whose closest HES episodes to diagnosis are an inpatient and outpatient record occurring on the same date. These are a special case of inpatient diagnosis.

Inpatient diagnosis (IP) – patients with a cancer diagnosis date related to a preceding inpatient HES episode (excluding patients already defined as special cases). An inpatient diagnosis is defined where the cancer diagnosis date is within the start and end of an episode. In addition, due to the potential for diagnosis to be confirmed following a relevant inpatient episode, a cancer diagnosis date that is within six months after the end of an episode and with no outpatient episode between would also be regarded as an inpatient diagnosis.

Outpatient diagnosis (OP) – patients with no inpatient HES episode preceding the cancer diagnosis date (as defined above) but with an outpatient HES attendance preceding the cancer diagnosis date.

Unknown (UN) – Unable to match cancer diagnosis date to any inpatient or outpatient HES episode. It is likely that, for these patients, the cancer diagnosis date was obtained from pathology records only, indicating diagnosis or treatment that only took place outside of a hospital setting (e.g. NHS patients seen in primary care, independent treatment centres or a community setting, and private patients seen and treated only in private hospitals).

Death Certificate Only diagnosis (DC) - The cancer registry receives a small number of cancer related death notifications, for which, despite extensive enquiries, they are unable to obtain additional information to register the disease details fully. This registration is regarded as Death Certificate Only (DCO) and the date of diagnosis is the same as that of the date of death.

2.7 Assigning the pathway group code

Each tumour was assigned a pathway group code based on the presence of inpatient and outpatient HES data as detailed in Table 2.1.

Table 2.1: Pathway Group codes

Pathway group	Description
A	Inpatient only in 6 months prior to diagnosis
B	Outpatient only in 6 months prior to diagnosis
C	Inpatient and outpatient in 6 months prior to diagnosis
D	No HES data 6 months prior to diagnosis
E	No HES data at all prior to diagnosis

2.8 Assigning the Route start-point

The start-point is determined by working backwards from the end-point as shown in Appendix 2. The characteristics of this start-point lead to a categorisation of Route:

- Routes that originated in an outpatient attendance use the outpatient source of referral of that attendance as the 'start-point' code;
- Routes that originated in an inpatient episode use the inpatient method of admission as the 'start-point' code;
- Routes where inpatient or outpatient data were unavailable the start-point codes may be assigned as null or unknown (this also includes DCOs).

A list of all possible 'start-point' codes is provided in Appendix 3.

2.9 Assigning the detailed Route to Diagnosis code

For each patient, a route end-point, the pathway group and the route start-point were derived and an overall detailed route code was defined by the concatenation of these three codes in the specific order: end-point–pathway group–start-point (e.g., IP-02-O03). This resulted in a total of 71 distinct routes to diagnosis codes, listed in Appendix 1.

2.10 Assigning the broad Route to Diagnosis category

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). Three routes reflect the urgency of referral (Emergency, Two-week Wait 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 below:

- **GP Referral**: includes routine and urgent referrals where the patient was not referred under the Two Week Wait referral route.
- **Two Week Wait**: urgent GP referrals with a suspicion of cancer.
- **Emergency Presentation**: an emergency route via A&E, emergency GP referral, emergency transfer, emergency admission or attendance.
- **Other Outpatient**: an elective route starting with an outpatient appointment that is either a self-referral, consultant to consultant referral, other or unknown referral (these referrals would not include patients originally referred under the Two Week Wait referral route).
- **Screen Detected**: flagged by the cancer registry as detected via the breast or cervical screening programmes.
- **Inpatient Elective**: where no earlier information can be found prior to admission from a waiting list, booked or planned.
- **DCO**: diagnosis by death certificate only.
- **Unknown**: no data available from IP or OP HES or from NCWT or screening.

The table in Appendix 1 was used to allocate Route categories from HES data.

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 prior to 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.

A case was linked to a CWT referral where a TWW had a decision to treat date within 62 days prior to 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 prior to 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 or the regional cancer registries respectively and no matching by date was performed.

3.0 Analytical techniques

This section details analytical methods used to interpret the outputs from the Routes to Diagnosis algorithm on the National Cancer Data Repository.

3.1 Tumour Grouping

For the analysis, 40 tumour types were identified, primarily based on their relevance to the NAEDI agenda. The list of tumour types by International Classification of Disease (ICD-10) codes is provided in Appendix 4.

3.2 Confidence intervals

Binomial confidence intervals for proportions of cancers diagnosed via a particular Route to Diagnosis are calculated using the Wilson score method[6]. Confidence intervals for survival analysis were calculated as part of the *strel* algorithm as defined below.

The figures (both proportions and survival) which are statistically significantly different at the 95% confidence interval limit are highlighted.

3.3 Funnel plots

Funnel plots for proportions of cancers diagnosed via a particular Route to Diagnosis were adapted from funnel plot templates[7] created by the Eastern Region Public Health Observatory.

3.4 Survival analysis

Relative survival is the ratio of the observed cumulative probability of survival in the study group and the survival that would have been expected if the group had only been subject to the background mortality in the general population (obtained from life tables). The particular life table used allowed for variations in background mortality by age, sex, region and social deprivation.

One-year relative survival was calculated using the *strel* tool developed by the London School of Hygiene and Tropical Medicine (LSHTM)[8] run within Stata version 10.

3.5 Colour shading

Tabular data is commonly presented with the background colour in each tabular cell related to the magnitude of the proportion in the cell. The extreme values of the background colours are set by the extreme values of the tabular data. Depending on the context of the table the extreme values might be those in the whole table or in one particular row or column. The colouring of each table should be considered a subjective 'guide to the eye' rather than having a fixed relationship to the magnitude of the data.

4.0 Data quality issues and limitations

This section outlines data quality issues in the raw data used and compares the frequency of cancer records with no information recorded ('Death Certificate Only' or DCO records) with that published by cancer registries themselves.

4.1 Screening data

An analysis of completeness of screening flags for England provided by cancer registries was undertaken, see Table 4.1. The breakdown by cancer registry shows a variation in the percentage of screen detected records assigned by each cancer registry, and in particular the figures for cervical in situ appear to be lower than expected for the majority of registries, see Table 4.2. This supported the exclusion of in situ codes from the main analysis.

Table 4.1: The number of records for England patients against each of the breast and cervical ICD-10 groupings (C50, C53, D05 and D06) by screen detected flag

Diagnosis Group		Screen Detected Flag			Percentage Screen Detected
		Yes	No	Total	
Breast	Malignant	31,039	79,134	110,173	28.2%
	In-Situ	7,902	5,642	13,544	58.3%
Cervix	Malignant	1,028	5,972	7,000	14.7%
	In-Situ	1,144	63,559	64,703	1.8%
Colorectal	Malignant	2086	89330	91416	2.3%
Total records		**43,199**	**243,637**	**286,836**	**15.1%**

Table 4.2: The percentage of records against each of the breast and cervical ICD-10 groupings (C50, C53, D05 and D06) by screen detected flag, broken down by English cancer registry Percentage of records that were Screen Detected

Percentage of records that were Screen Detected	Breast		Cervix		Colorectal
Cancer Registry	Malignant	In-Situ	Malignant	In-Situ	Malignant
Eastern Cancer Registration and Information Centre	29.8%	61.9%	0.5%	1.6%	2.6%
North West Cancer Intelligence Service	27.3%	55.6%	19.8%	0.2%	2.8%
Northern and Yorkshire Cancer Registry and Information Service	29.2%	64.4%	26.1%	0.7%	2.6%
Oxford Cancer Intelligence Unit	29.8%	58.4%	21.4%	0.0%	0.9%
South West Cancer Intelligence Service	28.7%	54.6%	21.1%	0.5%	2.1%
Thames Cancer Registry	25.2%	55.1%	0.2%	0.2%	1.6%
Trent Cancer Registry	30.0%	62.2%	9.5%	9.1%	2.6%
West Midlands Cancer Intelligence Unit	28.3%	58.6%	18.2%	1.6%	2.6%
Total records	**28.2%**	**58.3%**	**14.7%**	**1.8%**	2.3%

4.2 Death Certificate Only

Patients who were registered as a DCO on the National Cancer Data Repository and could not be matched to any of the data sources referenced in Section 2.2 above were assigned a DCO route grouping. However, there were patients registered as DCOs where additional information was found in inpatient and/or outpatient HES data which allowed these patients to be assigned a different route grouping, see Tables 4.3 and 4.4. This finding has important incidental implications for reducing the DCO rate for cancer registries, see Table 4.5. All tables below show the number of records as opposed to the number of distinct patients, which includes all records in the analysis (i.e. it does not exclude multiples).

Table 4.3: Comparison of the number of records assigned to the different routes to diagnosis against the number of records that have been flagged by the cancer registry as being DCO or non-DCO

Count of records		Registry DCO Flag		
Route to Diagnosis		**Yes**	**No**	**Total**
DCO		5,511	0	5,511
Not DCO	Emergency presentation	10,278	164,870	175,148
	GP referral	1,683	156,867	158,550
	Inpatient elective	452	42,489	42,941
	Other outpatient	1,322	69,982	71,304
	Screening	1	34,153	34,154
	TWW	109	192,251	192,360
	Unknown	72	59,627	59,699
Not DCO Total		13,917	720,229	734,146
Total records		**19,428**	**720,239**	**739,667**

Table 4.4: Comparison of percentage of records assigned to the DCO and non-DCO routes to diagnosis groupings against the percentage of records that have been flagged by the cancer registry as being DCO or non-DCO

Percentage of records	Registry DCO Flag		
Route to Diagnosis	**Yes**	**No**	**Total**
DCO	0.7%	0.0%	0.7%
Not DCO	1.9%	97.4%	99.3%
Total records	**2.6%**	**97.4%**	**100.0%**

Table 4.5: Comparison of the percentage of records assigned to the DCO routes to diagnosis groupings against the percentage of records that have been flagged by the cancer registry as being DCO, broken down by cancer registry

Percentage of Registry records Cancer Registry	Registry DCO Flag = Yes	Route to Diagnosis = DCO
Eastern Cancer Registration and Information Centre	0.3%	0.1%
North West Cancer Intelligence Service	5.4%	1.0%
Northern and Yorkshire Cancer Registry and Information Service	1.3%	0.6%
Oxford Cancer Intelligence Unit	1.6%	0.4%
South West Cancer Intelligence Service	1.7%	0.7%
Thames Cancer Registry	2.3%	1.1%
Trent Cancer Registry	1.8%	0.5%
West Midlands Cancer Intelligence Unit	6.2%	1.2%
Total records	**2.6%**	**0.7%**

4.3 Ethnicity

Reporting of ethnicity data in the 'Routes to Diagnosis' dataset is detailed in Table 4.6 and compared to the NCIN report "Cancer Incidence and Survival By Major Ethnic Group" which reported data from 2002-2006.

Table 4.8: Comparison of ethnicity recording in the NCIN report "Cancer Incidence and Survival By Major Ethnic Group" against the Routes to diagnosis dataset.

Ethnicity data profile	NCIN Ethnicity report		Routes to Diagnosis	
	People	Percentage	People	Percentage
White	435,168	73%	571,293	79.2%
Asian	6,685	1%	10,331	1.4%
Black	6,540	1%	8,152	1.1%
Chinese	651	0%	1,150	0.2%
Mixed	1,058	0%	1,972	0.3%
Other ethnic group	3,194	1%	4,313	0.6%
Unknown	145,299	24%	124,352	17.2%
Total	**598,595**	**100%**	**721,563**	**100%**

5.0 Further methodological development

This section records points noted in the development of the Routes to Diagnosis algorithm to further improve or develop it.

5.1 Outstanding issues within the Routes to Diagnosis algorithm

Several minor issues were noted during the development of the Routes to Diagnosis algorithm for the 2nd iteration. These are described below with steps that might be taken to resolve them.

DNA and cancelled status of outpatient episode

A small percentage of outpatient episodes, while present in the dataset, are coded as "DNA" (indicating that the patient Did Not Attend) or cancelled. The project team decided not to remove these episodes in the belief that information contained in the episode might still be relevant to the patient's Route to Diagnosis. This should be reviewed as part of further Routes to Diagnosis development work.

Multiple outpatient attendances on same day

A small percentage of outpatients attendances occur on the same day as another outpatient appointment. In these cases the temporal order was assigned randomly for purposes of deciding which was closer to the time of diagnosis. This should be reviewed as part of further Routes to Diagnosis development work and also whether any fields exist within the outpatient dataset that could be used to assign priority to one outpatient attendance.

5.2 Expansion of data sources of algorithm

The Routes to Diagnosis algorithm relies on Cancer Registration data, plus in- and out-patient HES data, Cancer Waiting Times data, and data from the Breast, Cervical (via the cancer registries) and Colorectal screening services. Including further data sources may add to the robustness or utility of the algorithm.

Expand to include Accident and Emergency data

A&E HES data may provide more complete information on Emergency Presentations or enable them to be analysed at a more granular scale. The feasibility of building A&E HES data into the algorithm should be explored.

Expand to include Diagnostic Imaging Dataset (DID)

Diagnostic imaging carried out in secondary care should be picked up by the Routes to Diagnosis algorithm as part of an outpatient attendance or inpatient episode. However, imaging conducted in primary care will not currently be captured. As DID data becomes available the feasibility of building it into the algorithm should be explored.

Expand to include Primary care data

The secondary care setting is the focus of the datasets currently used by the algorithm (except screening). Adding primary care data would allow the parts of the Route to Diagnosis which takes place in Primary Care to be mapped. While there are not presently any primary care datasets which have complete national coverage the feasibility of including primary care data in the algorithm should be explored.

References

1. Routes to Diagnosis 2012 available online at *www.ncin.org.uk/*

2. *Routes to Diagnosis for cancer - Determining the patient journey using multiple routine datasets (2012). L Elliss-Brookes et al. British Journal of Cancer Vol 107, No 8.*

3. *NCIN Data Briefing, Routes to Diagnosis available online at http://www.ncin.org.uk/publications/data_briefings/routes_to_diagnosis.aspx*

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

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

6. APHO Technical Briefing 3: Commonly used public health statistics and their confidence intervals. Available online at: www.apho.org.uk/resource/view.aspx?RID=48457

7. Analytical Tools for Public Health: Funnel plot for proportions and percentages http://www.apho.org.uk/resource/view.aspx?RID=47241

8. 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.htm on 8th February 2010 Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, UK

Authorship and Acknowledgements

The Routes to Diagnosis 2nd iteration project team consisted of Lucy Elliss-Brookes, Alex Ives, Matt Greenslade, Sean McPhail, Mike Richards, and Jon Shelton.

The project team is grateful for the help and support of Chris Carrigan, the late Brian Cottier, Sara Hiom, Tariq Malik, Jodie Moffat, Eva Morris, Andy Pring, James Thomas, Catherine Thomson, Julia Verne and the other staff of the National Cancer Intelligence Network, English Cancer Registries, Cancer Research UK, and the National Cancer Services Analysis Team.

Version Control

The document is version 1.0, September 2012.

Glossary

Pathway Group

A classification that is created for each tumour according to the presence or absence of inpatient and outpatient HES data in the 6 months prior to diagnosis

Route to Diagnosis

A 'Route to Diagnosis' is defined as the sequence of interactions between the patient and the healthcare system which lead to a diagnosis of cancer, based on the end point, the pathway and the referral route into secondary care. Depending on context it might either be a 'detailed' route, e.g. IP-C-O4, or a broad summary route, e.g. "Emergency Presentation".

Route start-point

The start point is the first recorded clinical care event that the Route to Diagnosis Algorithm

Route end-point

The end-point was assumed to be the clinical care event that led most immediately to diagnosis.

Appendix 1: Routes to diagnosis codes

A list of all 'Routes to Diagnosis' codes is provided in Table A1.1. The route code is in the form of route end point – pathway group– start-point.

Table A1.1: Route to Diagnosis codes

Number	Route Code	Route Group
1	DC-D-DCO	DCO
2	DC-E-DCO	DCO
3	IP-A-I11	Inpatient elective
4	IP-A-I12	Inpatient elective
5	IP-A-I13	Inpatient elective
6	IP-A-I21	Emergency presentation
7	IP-A-I22	Emergency presentation
8	IP-A-I23	Emergency presentation
9	IP-A-I24	Emergency presentation
10	IP-A-I31	Inpatient elective
11	IP-A-I32	Inpatient elective
12	IP-A-I81	Inpatient elective
13	IP-A-I82	Inpatient elective
14	IP-A-I83	Inpatient Elective
15	IP-A-I84	Inpatient Elective
16	IP-A-I89	Inpatient Elective
17	IP-A-I99	Unknown
18	IP-A-UNK	Unknown
19	IP-C-O01	Emergency presentation
20	IP-C-O02	Other outpatient
21	IP-C-O03	GP referral
22	IP-C-O04	Emergency presentation
23	IP-C-O05	Other outpatient
24	IP-C-O06	Other outpatient
25	IP-C-O07	Other outpatient
26	IP-C-O08	Other outpatient
27	IP-C-O10	Emergency presentation
28	IP-C-O11	Other outpatient
29	IP-C-O12	GP referral
30	IP-C-O13	Other outpatient
31	IP-C-O17	Screening
32	IP-C-O92	Other outpatient
33	IP-C-O93	Other outpatient
34	IP-C-O97	Other outpatient
35	OP-B-O01	Emergency presentation
36	OP-B-O02	Other outpatient
37	OP-B-O03	GP referral
38	OP-B-O04	Emergency presentation
39	OP-B-O05	Other outpatient
40	OP-B-O06	Other outpatient
41	OP-B-O07	Other outpatient
42	OP-B-O08	Other outpatient
43	OP-B-O10	Emergency presentation
44	OP-B-O11	Other outpatient
45	OP-B-O12	GP referral
46	OP-B-O13	Other outpatient
47	OP-B-O17	Screening
48	OP-B-O92	Other outpatient
49	OP-B-O93	Other outpatient
50	OP-B-O97	Other outpatient
51	OP-B-O99	Unknown
52	SC-C-null	Unknown
53	SC-C-O01	Emergency presentation
54	SC-C-O02	Other outpatient
55	SC-C-O03	GP referral
56	SC-C-O04	Emergency presentation
57	SC-C-O05	Other outpatient
58	SC-C-O06	Other outpatient
59	SC-C-O07	Other outpatient
60	SC-C-O08	Other outpatient
61	SC-C-O10	Emergency presentation
62	SC-C-O11	Other outpatient
63	SC-C-O12	GP referral
64	SC-C-O13	Other outpatient
65	SC-C-O17	Screening
66	SC-C-O92	Other outpatient
67	SC-C-O93	Other outpatient
68	SC-C-O97	Other outpatient
69	SC-C-O99	Unknown
70	UN-D-UNK	Unknown
71	UN-E-UNK	Unknown

Appendix 2: Algorithmic flow diagrams

Figure A2.1: Flow diagram for allocating the end point of the route using inpatient and outpatient HES data.

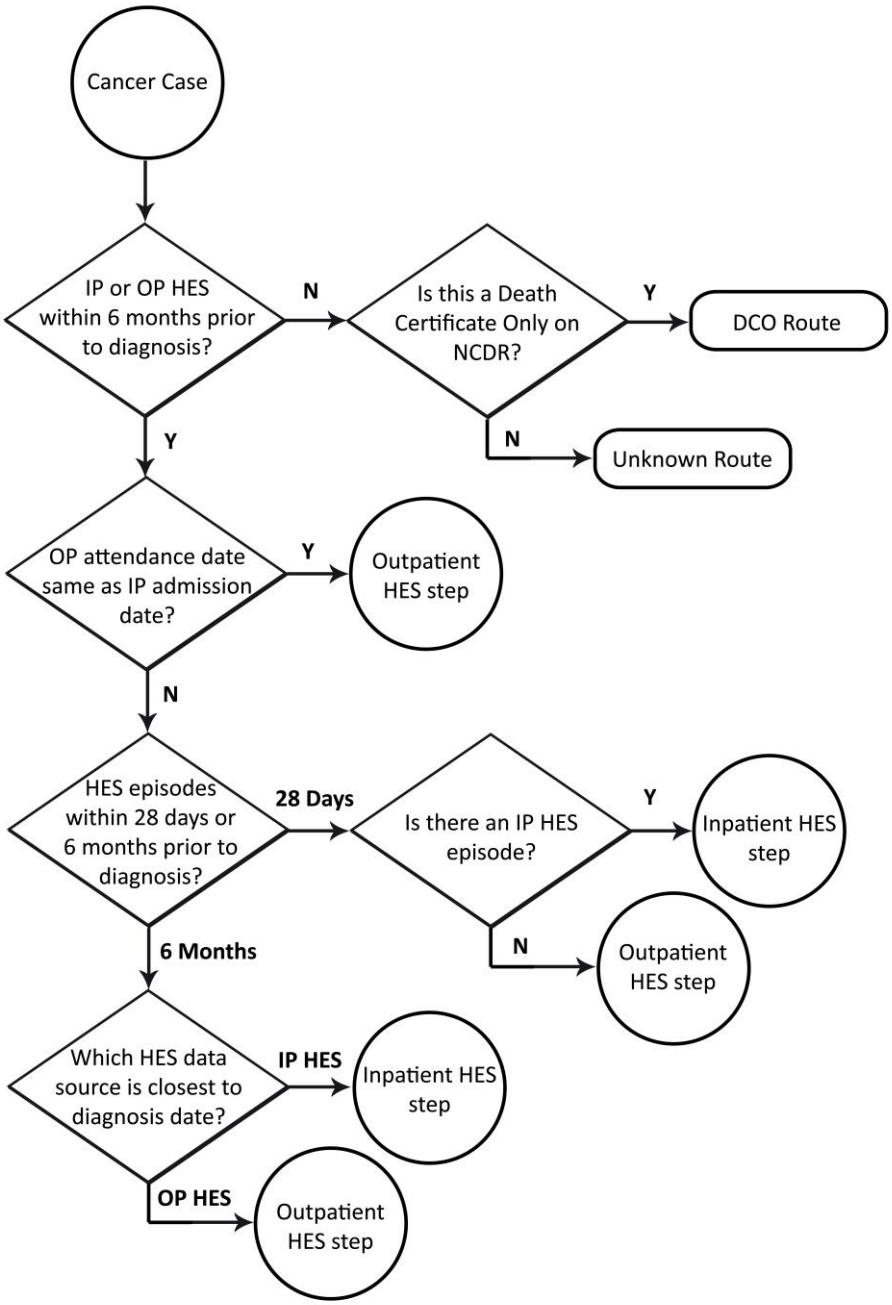


Figure A2.2: Flow diagram for finding the start point or prior step for an inpatient step in a route.

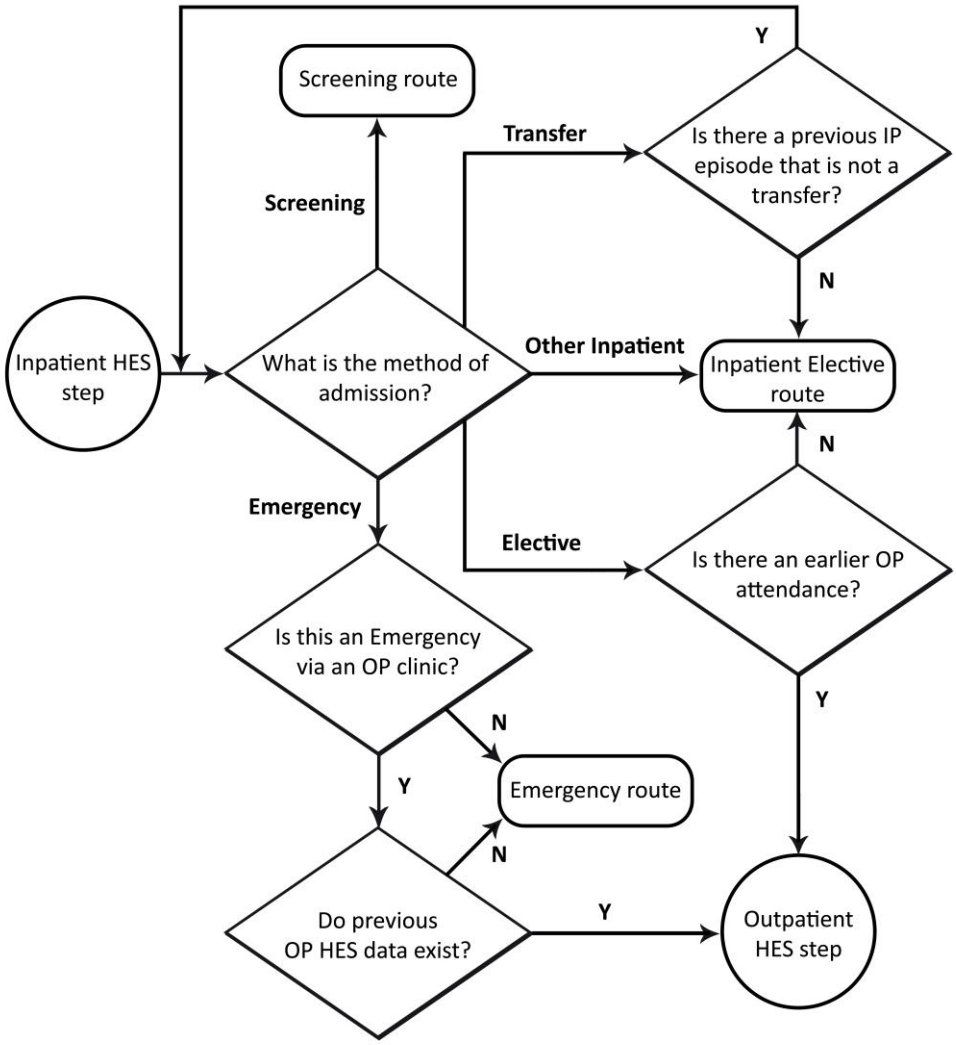
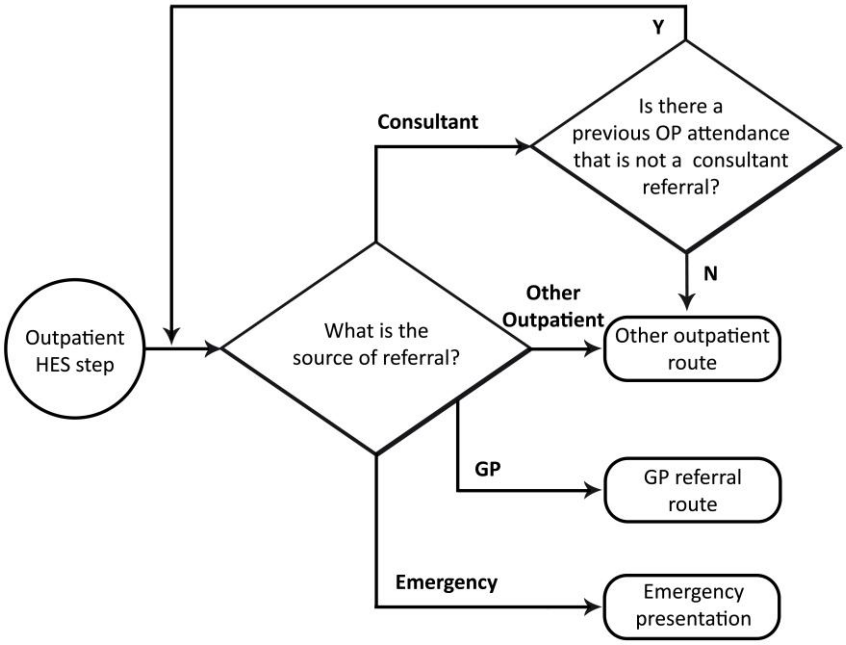


Figure A2.3: Flow diagram for finding the start point or prior step for an outpatient step in a route.



Appendix 3: Start-point codes

A list of all 'start-point' codes is provided in Table A3.1. Codes that commence with an 'I' indicates an inpatient method of admission while an 'O' indicates an outpatient source of referral.

Table A3.1: Start-point codes

Start point code	Start point Description
DCO	DCO
I11	Elective: from waiting list
I12	Elective: booked
I12	Elective:booked
I21	Emergency: via Accident and Emergency (A&E) services, including the casualty department of the provider
I22	Emergency: via general practitioner (GP)
I23	Emergency: via Bed Bureau, including the Central Bureau
I24	Emergency: via consultant outpatient clinic
I28	Emergency: other means, including patients who arrive via the A&E department of another healthcare provider
I31	Maternity: where the baby was delivered after the mothers admission
I32	Maternity: where the baby was delivered before the mothers admission
I81	Transfer of any admitted patient from another hospital provider other than in an emergency; this does not include admissions to high security psychiatric hospitals (HSPH)
I82	Other: babies born in health care provider
I83	Other: babies born outside the health care provider, except when born at home is intended
I98	Not applicable (eg other maternity event)
I99	Not known
O01	Following an emergency admission
O02	Following a domiciliary visit
O03	Referral from a general medical practitioner
O04	Referral from an accident and emergency department
O05	Referral from a consultant, other than in an accident and emergency department
O06	Self referral
O07	Referral from prosthetist
O08	Other source of referral
O10	Following an accident and emergency attendance
O11	Other
O12	Referral from GP with special interest
O13	Referral from a specialist nurse (secondary care)
O14	Referral from an allied health professional
O15	Referral from an optometrist
O16	Referral from an orthopist
O17	Referral from a national screening programme
O92	General dental practitioner
O93	Community dental service
O97	Other - not initiated by the consultant responsible for the consultant outpatient episode
O99	Not known
UNK	Unknown

Appendix 4: Tumour categories

Table A4.1: Tumour categories with associated ICD-10 codes.

ICD-10 Code	Cancer Site
C00-C97 excl. C44	All Cancers
C00, C05, C07, C08, C11-C14	Head and neck - Other (excl. oral cavity, oropharynx and thyroid)
C00, C05, C11, C14	Head and neck - Other sites of the lip, oral cavity and pharynx
C01, C09-C10	Head and neck - Oropharynx
C02-C04, C06	Head and neck - Oral cavity
C07-C08	Head and neck - Salivary glands
C12-C13	Head and neck - Hypopharynx
C15	Oesophagus
C16	Stomach
C18-C20	Colorectal
C22	Liver
C25	Pancreas
C32	Head and neck – larynx
C33-C34	Lung
C40-C41	Sarcoma: bone
C40-C41, C46, C48	Sarcoma: other
C43	Melanoma
C45	Mesothelioma
C48	Sarcoma: retroperitoneum and peritoneum
C49	Sarcoma: connective and soft tissue
C50	Breast
C51	Vulva
C53	Cervix
C54-C55	Uterus
C56	Ovary
C61	Prostate
C62	Testis
C64-C66, C68	Kidney and unspecified urinary organs
C67	Bladder
C70-C72	Central Nervous System
C73	Head and neck – thyroid
C81	Hodgkin lymphoma
C82-C85	Non-Hodgkin lymphoma
C88-C90	Multiple myeloma
C910	Leukaemia: acute lymphoblastic
C911	Leukaemia: chronic lymphocytic (CLL)
C91-C95 exc. C910, C911, C921, C920, C924-C925, C930, C940, C942	Leukaemia: rarer types
C91-C95 exc. C911, C920, C924-C925, C930, C940, C942	Leukaemia: other (all excluding AML and CLL)
C920, C924-C925, C930, C940, C942	Leukaemia: acute myeloid (AML)
C921	Leukaemia: Chronic myeloid

National Cancer Intelligence Network
18th Floor, Portland House
Bressenden Place
London, SW1E 5RS

enquiries@ncin.org.uk

www.ncin.org.uk

NCIN
national cancer
intelligence network
Using information to improve quality & choice