Data quality and completeness report:

Upper Gastrointestinal Site Specific Clinical Reference Group (SSCRG)

1. Introduction

The National Cancer Intelligence Network (NCIN) Upper Gastrointestinal Cancer Site Specific Clinical Reference Group covers oesophago-gastric (OG) cancers (including oesophageal and stomach cancer) and primary hepatic, pancreatic and biliary cancers (including primary liver, biliary, ampulla of Vater, duodenum, gallbladder and pancreas), (Appendix 1). Thames Cancer Registry investigates these cancers using data from the National Cancer Repository dataset (NCRD). The NCRD contains information from each of the eight English cancer registries on all patients diagnosed with cancer in their catchment areas and includes any relevant treatment information in the six months following diagnosis from the Hospital Episode Statistics (HES) dataset. HES data is supplied to the English cancer registries by the NHS Information Centre.

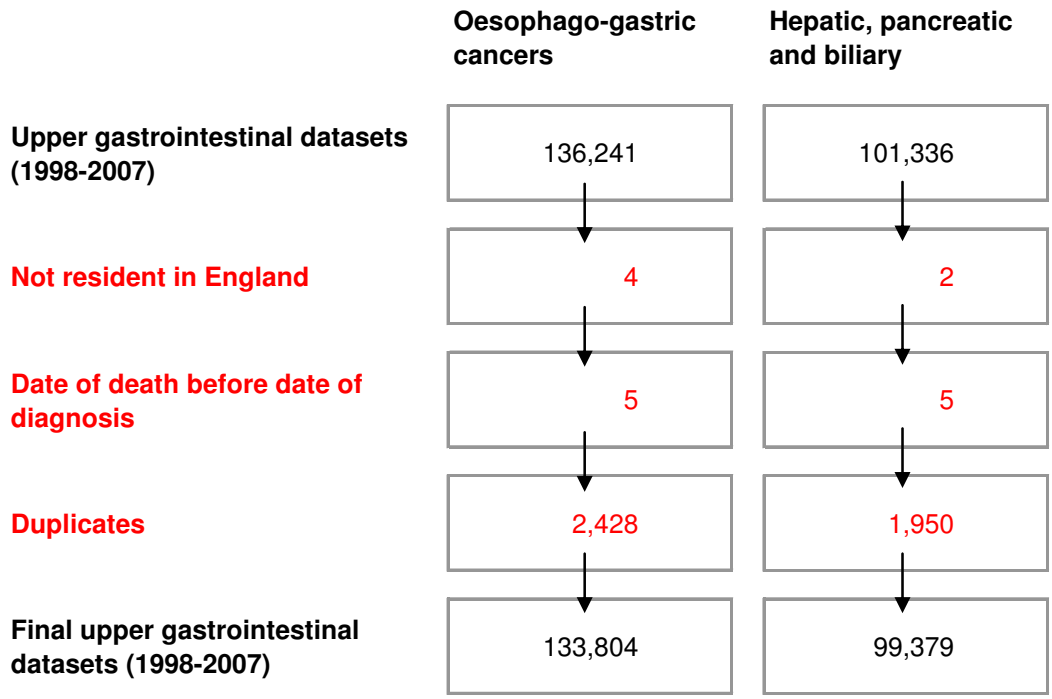
It is important to analyse the quality of the data as large proportions of missing or poor quality information will lead to potentially inaccurate conclusions being drawn. It will also mean that some more detailed analysis on specific sub groups would be difficult. It is vital to record the quality of the data to ensure improvements can be made if found to be necessary. An annual report will help drive and measure any improvements.

This report aims to explore the data quality and completeness of the upper gastrointestinal cancer dataset. It reports on data on patients diagnosed between 1998 and 2007.

2. Methods

Data were extracted from the National Cancer Repository Dataset on all patients diagnosed with upper gastrointestinal cancer between 1998 and 2007. The initial datasets consisted of 136,241 patients diagnosed with oesophago-gastric cancer (OG) and 101,336 patients diagnosed with hepatic, pancreatic and biliary cancers (HPB). A small number of patients were excluded as they were not resident in England (n=6) or their date of death was before their date of diagnosis (n=10) (Figure 1). 2,428 (1.8%) OG and 1,950 (1.9%) HPB duplicates were also removed. Therefore, the final dataset consisted of data on 133,804 patients diagnosed with OG cancer and 99,379 patients diagnosed with primary HPB cancers.

Figure 1: Patient flow within the oesophago-gastric (OG) and primary hepatic, pancreatic and biliary (HPB) cancer datasets.



2.1 Data quality

The quality of the dataset was investigated for the main cancer sites including cancers of the oesophagus (ICD10 C15), stomach (ICD10 C16), duodenum (ICD10 C17.0), primary liver (ICD10 C22), gallbladder (ICD10 C23), biliary (ICD10 C24) and pancreas (ICD10 C25), (see appendix 1).

Data were analysed at cancer registry level (Table 1). The graphs and accompanying text will refer to each registry by their code.

Table 1: List of the eight English cancer registries.

Cancer registry code	Cancer registry name
ECRIC	Eastern Cancer Registration and Information Centre
NWCIS	North West Cancer Intelligence Service
NYCRIS	Northern & Yorkshire Cancer Registry and Information Service
Oxford	Oxford Cancer Intelligence Unit
SWCIS	South West Cancer Intelligence Service
Thames	Thames Cancer Registry
Trent	Trent Cancer Registry
WMCIU	West Midlands Cancer Intelligence Unit

The data quality measures investigated are listed below:

a) Proportion of death certificate only registrations (DCO)

Many registrations for rapidly fatal cancers are initiated by the patient's death certificate. These registrations are followed up in hospital systems or in the HES dataset. Many cases are found and their details are updated to form a complete registration. Those that are not found remain death certificate only registrations (DCOs). These registrations have limited information and their date of diagnosis is the same as their date of death. They therefore have to be excluded from some analyses.

b) Proportion of patients with an unspecified anatomical site

The proportions of patients with an unspecified anatomical site were calculated. This included patients with an International Classification of Diseases version 10 (ICD10) 4 digit code of Cxx.8 (overlapping lesion of [specific] cancer) and Cxx.9 ([specific] cancer, unspecified). See Appendix 2 for full list of codes. Large proportions of patients with an unspecified anatomical site will limit our ability to analyse these cancers by specific subgroups.

c) Proportion of patients by basis of diagnosis

The proportions of patients by their basis of diagnosis were calculated. This included microscopically verified, any other test (e.g. Computed Tomography (CT) scan, X-ray), not known or missing. Cases that are not microscopically verified will not have a valid morphology.

d) Proportion of patients with a missing ethnicity

Ethnicity has historically been poorly recorded in cancer registry datasets. Since 1995 it has been mandatory to collect ethnicity information within hospitals and therefore the NCRD includes ethnicity from the hospital episode statistics (HES) dataset. Large proportions of patients with a missing ethnicity code will make studies focussing on ethnicity less robust.

e) Proportion of patients with a missing stage

Stage is an important indicator of the prognosis and will influence the treatment that patients receive. This report shows the proportions of patients that had enough information recorded in the dataset to allow a stage to be derived. Staging information was considered to be available if a patient had either a record of metastasis, any pathological TNM information, any clinical TNM information or an original stage (in that order). Metastases were recorded as "yes", "no" or "not known". The T, N and M fields were considered separately and were included if they held a valid code. If the fields were blank or contained an "X" these were assumed to be zero. The original stage was the stage supplied by each cancer registry. The proportion of patients with a missing stage was also calculated by year of diagnosis.

f) Proportion of patients with no linked HES records

The proportion of patients with no linked HES records were calculated for each cancer registry. No linked HES records could indicate that the matching has not been successful for that patient and as a result their treatment information may not have been included in our dataset. Also, the subset of HES data received by the cancer registries only includes patients with a diagnosis of cancer. Patients may have had surgery for their cancer, but have no cancer diagnosis in HES. Therefore, their surgery would not be linked to their cancer registration record. However, it could also mean that the patient has had no inpatient activity. This will be important to consider in any future treatment analysis.

2.2 Completeness

The completeness of the cancer registry dataset has often been questioned. It is important to ascertain an estimate of how many potential cancer registrations are missed each year. Large proportions of missing registrations could affect survival analyses with estimates being too low if patients with better prognoses are missed.

Using the Hospital Episode Statistics database, patients who had a diagnosis of cancer between 1998 and 2007 and who had no matching record in the cancer registry dataset were identified (HES-onlys). HES-only registrations were then narrowed down to only include those with a relevant procedure code related to the cancer in question (see Appendix 3). The combination of diagnosis and surgery codes taken together increases the certainty that these patients are true cancer cases, rather than just a record of a suspicion of cancer. These registrations are considered most likely to have been missed by the cancer registration process. This analysis was carried out at a patient level.

HES-only registrations were considered alongside the cancer registration records and an incompleteness measure was calculated. This was stratified by sex, age, year of diagnosis and cancer registry using the same method recently employed by Møller and colleagues (2010).

2.3 Cancer sub groups

The data quality analysis was also run for each of the twelve sub groups defined by the Site Specific Clinical Reference Group for more detailed analysis in future reports (see Appendix 1). The OG cancers sub groups include cancers of the upper and middle oesophagus, lower oesophagus, oesophagus with an unspecified anatomical site, cardia, distal stomach and stomach with an unspecified anatomical site. The HPB cancers sub groups include cancers of the duodenum, liver, biliary, gallbladder, ampulla of Vater and pancreas.

The tables show the proportion of death certificate only registrations between the period 1998 and 2007. They also show the proportion of patients with a missing ethnicity, a missing stage and with no linked record in HES. This part of the analysis excludes death certificate only registrations.

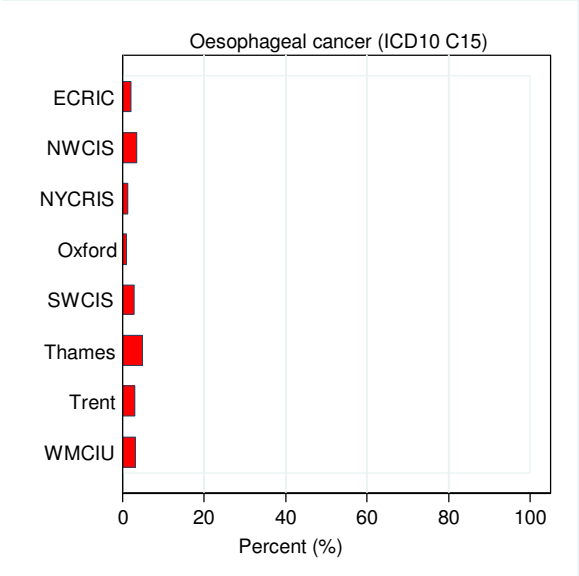
3. Results

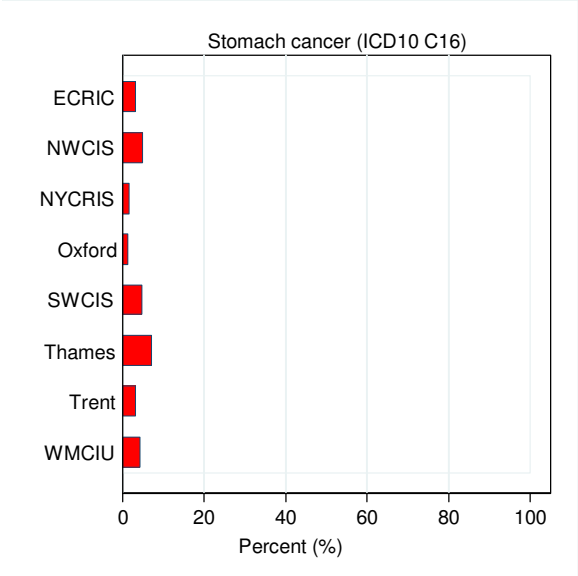
3.1 Quality of the upper gastrointestinal cancer dataset, England, 1998-2007

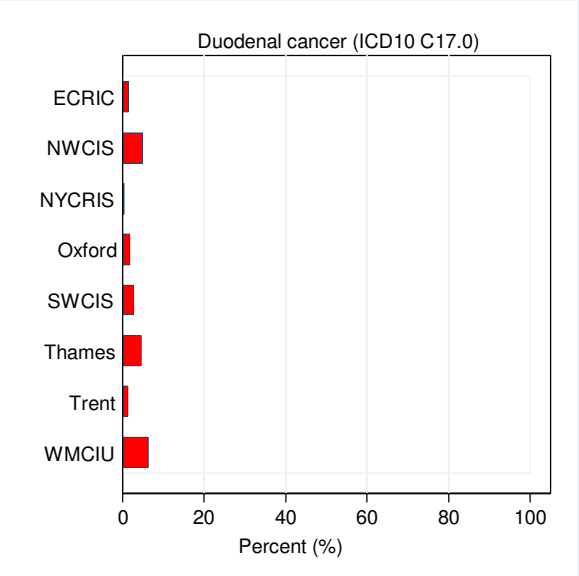
	Oesophageal cancer (ICD10 C15)		Stomach cancer (ICD10 C16)		Duodenal cancer (ICD10 C17.0)		Liver cancer (ICD10 C22)		Gallbladder cancer (ICD10 C23)		Biliary cancer (ICD10 C24)		Pancreatic cancer (ICD10 C250)	
	61,875		71,929		2,684		23,269		4,550		6,566		62,310	
Death certicate onlys														
Death certicate only	1,930	(3.1)	3,060	(4.3)	87	(3.2)	2,464	(10.6)	379	(8.3)	270	(4.1)	6,177	(9.9)
Non-DCO registrations	59,945	(96.9)	68,869	(95.7)	2,597	(96.8)	20,805	(89.4)	4,171	(91.7)	6,296	(95.9)	56,133	(90.1)
Anatomical site														
No anatomical subgroup	32,285	(53.9)	34,969	(50.8)	-	-	-	-	-	-	638	(10.1)	26,104	(46.5)
Known anatomical site	27,660	(46.1)	33,900	(49.2)	-	-	-	-	-	-	5,658	(89.9)	30,029	(53.5)
Basis of diagnosis														
Microscopically verified	54,517	(90.9)	62,412	(90.6)	2,290	(88.2)	10,012	(48.1)	2,910	(69.8)	4,411	(70.1)	24,734	(44.1)
Any other test	4,819	(8.0)	5,757	(8.4)	281	(10.8)	10,164	(48.9)	1,173	(28.1)	1,793	(28.5)	29,428	(52.4)
Not known	474	(0.8)	534	(0.8)	22	(0.8)	433	(2.1)	58	(1.4)	64	(1.0)	1,105	(2.0)
Missing	135	(0.2)	166	(0.2)	4	(0.2)	196	(0.9)	30	(0.7)	28	(0.4)	866	(1.5)
Ethnicity														
Known	49,051	(81.8)	54,605	(79.3)	2,114	(81.4)	15,583	(74.9)	3,047	(73.1)	5,045	(80.1)	41,180	(73.4)
Not known	10,894	(18.2)	14,264	(20.7)	483	(18.6)	5,222	(25.1)	1,124	(26.9)	1,251	(19.9)	14,953	(26.6)
Stage														
Known	11,877	(19.8)	15,683	(22.8)	498	(19.2)	2,187	(10.5)	1,133	(27.2)	1,072	(17.0)	13,258	(23.6)
Not known	48,068	(80.2)	53,186	(77.2)	2,099	(80.8)	18,618	(89.5)	3,038	(72.8)	5,224	(83.0)	42,875	(76.4)
No linked record in Hospital Episode Statistics														
Link	56,947	(95.0)	64,156	(93.2)	2,416	(93.0)	18,170	(87.3)	3,561	(85.4)	5,801	(92.1)	49,887	(88.9)
No linked	2,998	(5.0)	4,713	(6.8)	181	(7.0)	2,635	(12.7)	610	(14.6)	495	(7.9)	6,246	(11.1)

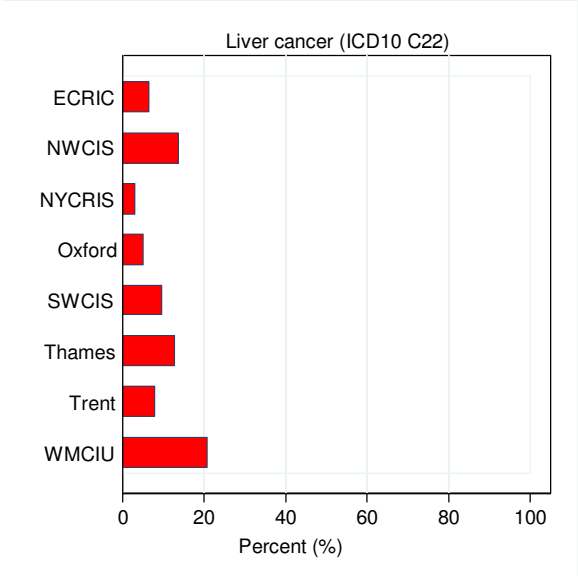
3.2 Proportion of death certificate only registrations by cancer registry

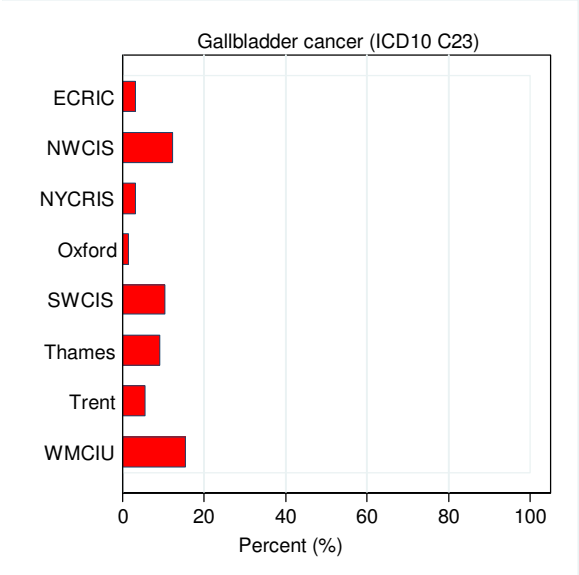
The following graphs show the proportion of death certificate only registrations over the period 1998 to 2007.

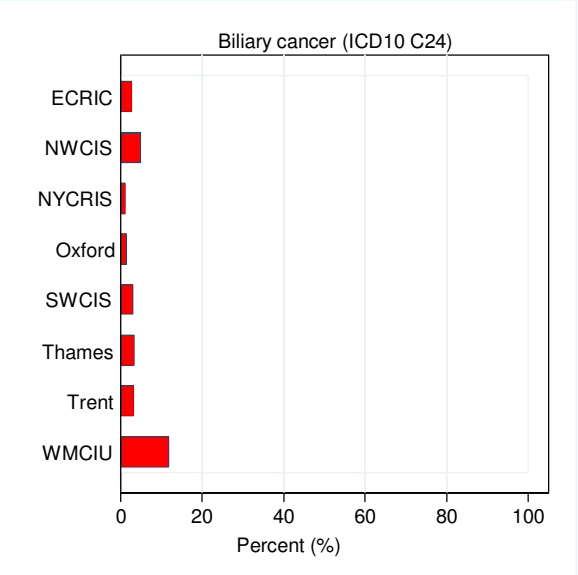


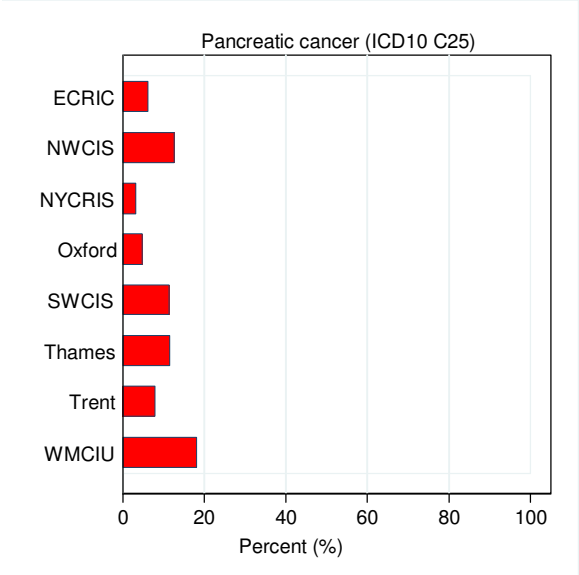












Oesophagus: DCOs ranged from 1.0% in Oxford to 5.0% in Thames.

Stomach: DCOs ranged from 1.4% in Oxford to 7.3% in Thames.

Duodenum: DCOs ranged from 0.5% in NYCRIS to 6.4% in WMCIU.

Liver: DCOs ranged from 3.1% in NYCRIS to 20.8% in WMCIU.

Gallbladder: DCOs ranged from 1.6% in Oxford to 15.5% in WMCIU.

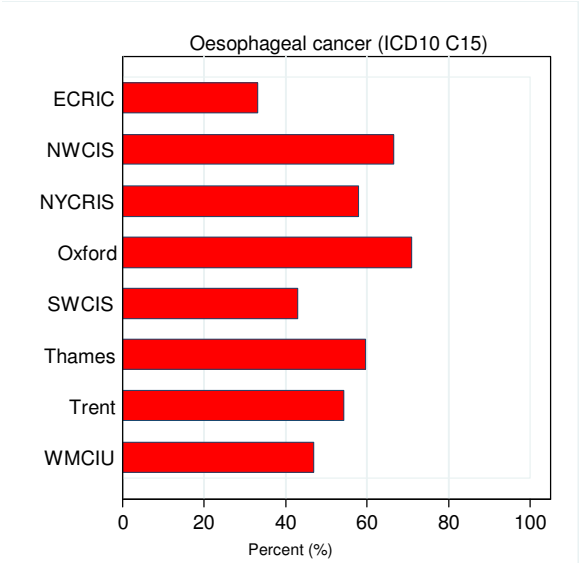
Biliary: DCOs ranged from 1.2% in NYCRIS to 11.9% in WMCIU.

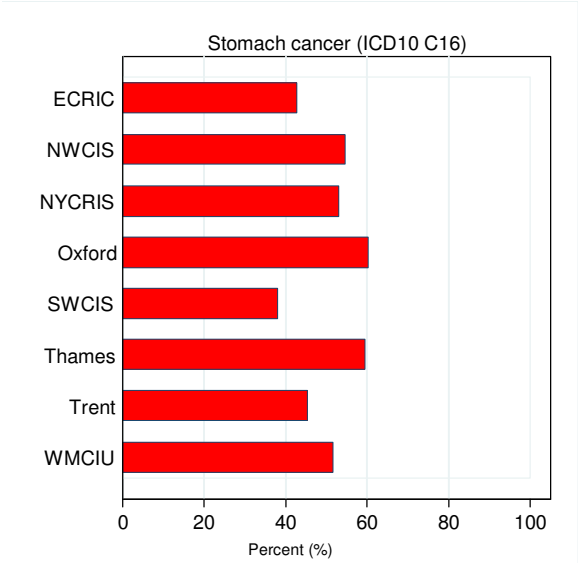
Pancreas: DCOs ranged from 3.2% in NYCRIS to 18.2% in WMCIU.

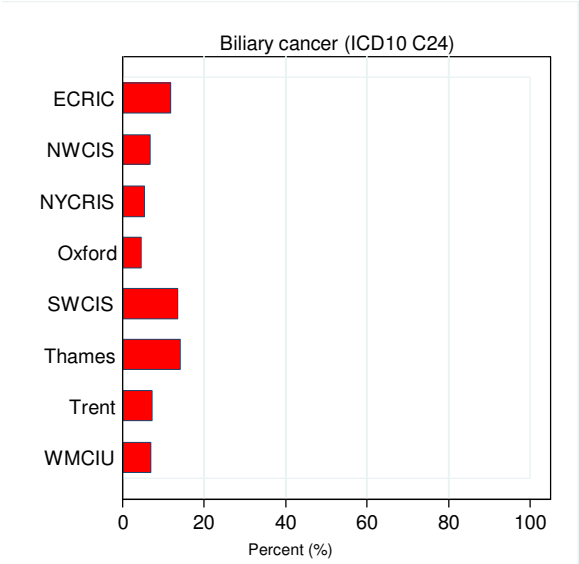
The proportion of death certificate only registrations ranged from 0.5% to 20.8%, although typically remained below 10%. Primary liver, gallbladder and pancreatic cancer had higher proportions of DCO registrations.

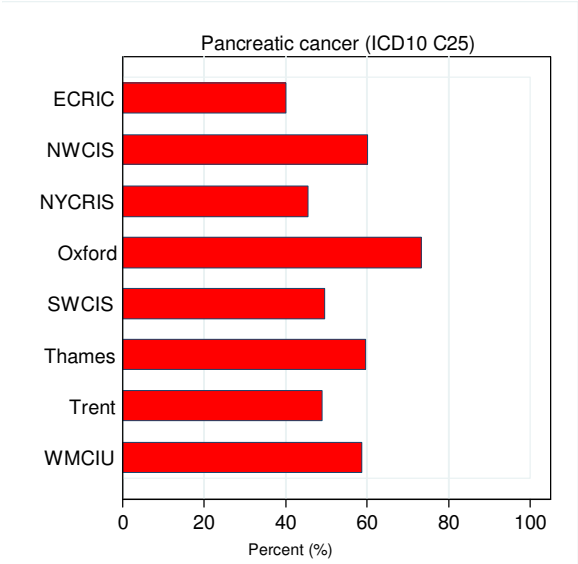
3.3 Proportion of patients with an unspecified anatomical site by cancer registry

The following graphs show the proportion of patients with an unspecified anatomical site over the period 1998 and 2007. This analysis excludes death certificate only registrations.







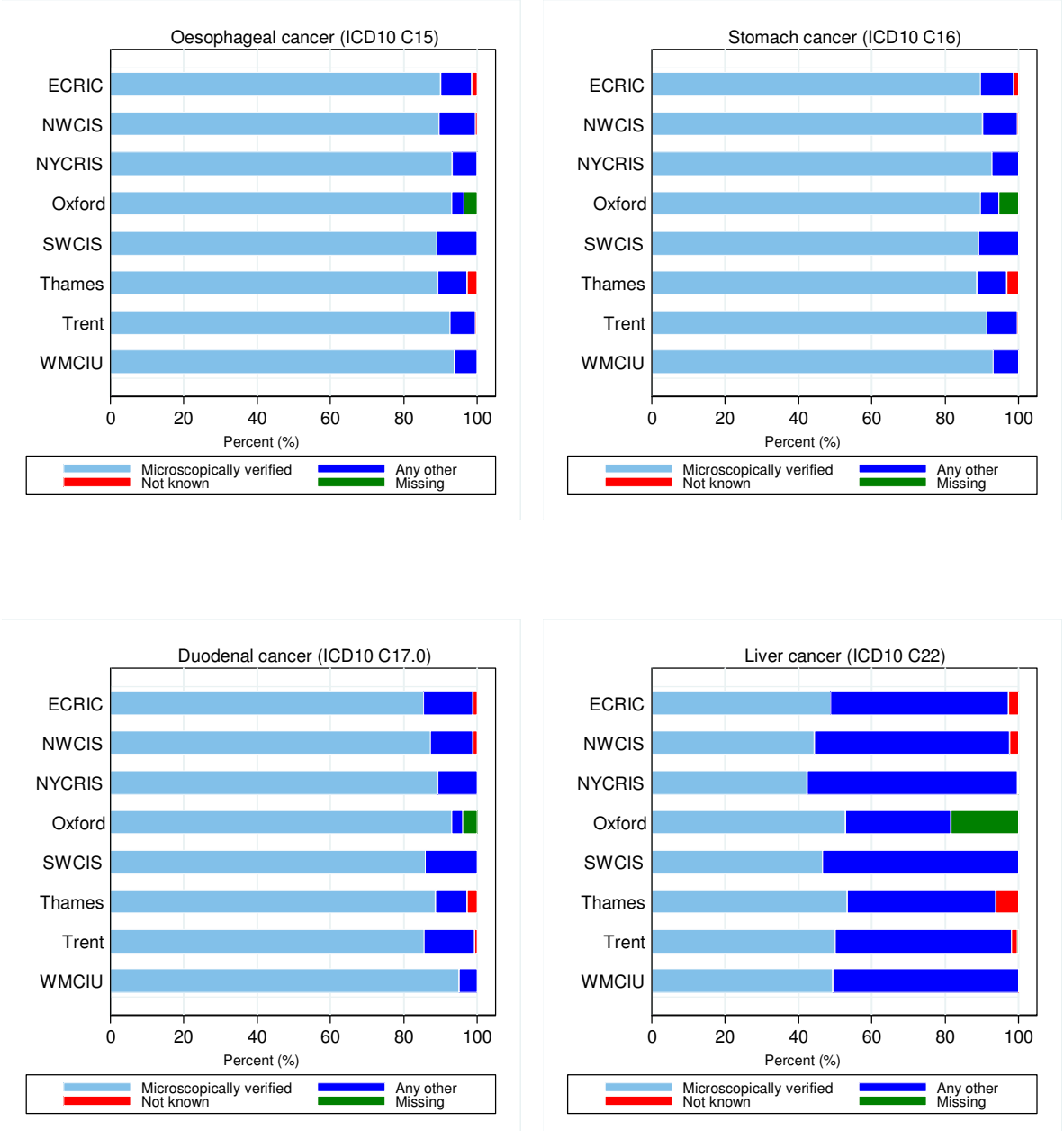


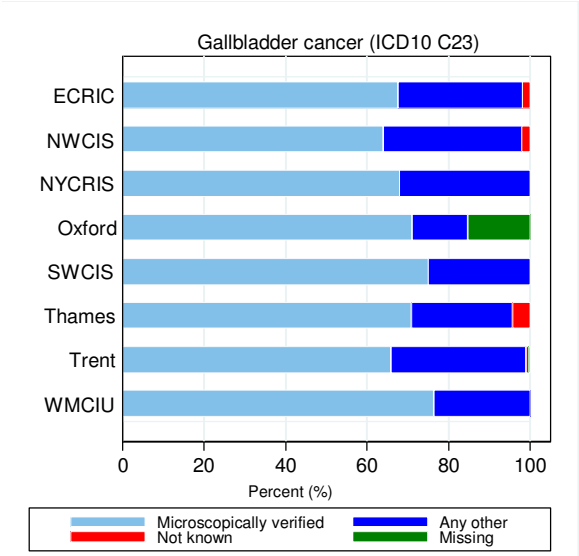
The proportions of patients with an unspecified anatomical site was higher in patients with oesophageal, stomach and pancreatic cancer. For oesophageal cancer, ECRIC (33.4%) had the least patients with an unspecified anatomical site and Oxford (71.2%) the most. Stomach cancer with an unspecified site ranged from 38.0% in NYCRIS and 60.5% in Oxford. Pancreatic cancer had 40.1% with an unspecified site in ECRIC and 73.3% in Oxford. Biliary cancers had less than 15% of registrations with an unspecified site.

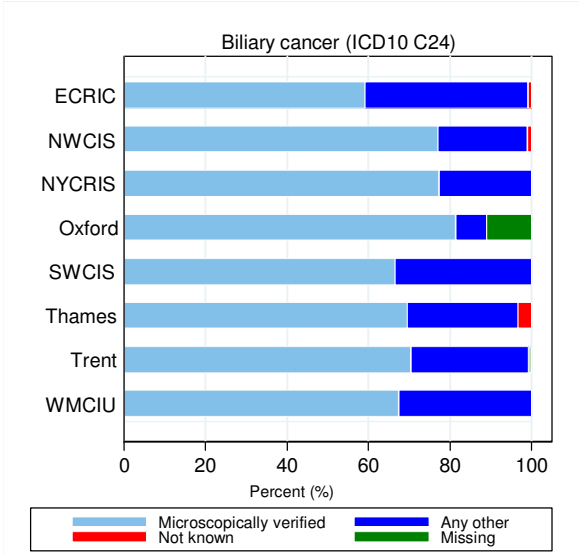
Duodenal cancer is defined by the ICD10 4 digit code of C17.0 (see appendix 2). Those with an unspecified anatomical location in the C17 (malignant neoplasm of the small intestine) group are defined as C17.8 (overlapping lesion of small intestine) and C17.9 (small intestine, unspecified). In addition to cancers of the duodenum these codes will also include cancers of the jejunum, ileum and Meckel's diverticulum, all of which are not included under the upper gastrointestinal site specific clinical reference group. Therefore, the proportions of cases with an unspecified site for duodenal cancer were not included in this report. Also, gallbladder cancers are coded as ICD10 C23. There are no further divisions in this group and consequently no unspecified anatomical locations.

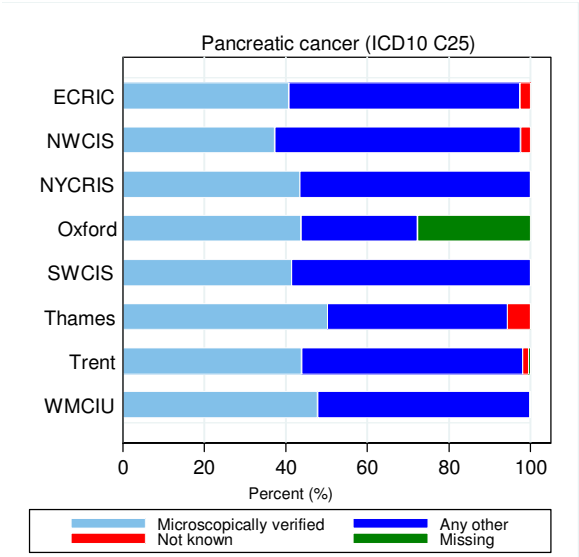
3.4 Proportion of patients by basis of diagnosis by cancer registry

The following graphs show the proportion of patients by each basis of diagnosis category for the period 1998 to 2007. This analysis excludes death certificate only registrations.









The proprtions of microscopically verified cases ranged from 37.4% to 95.1%.

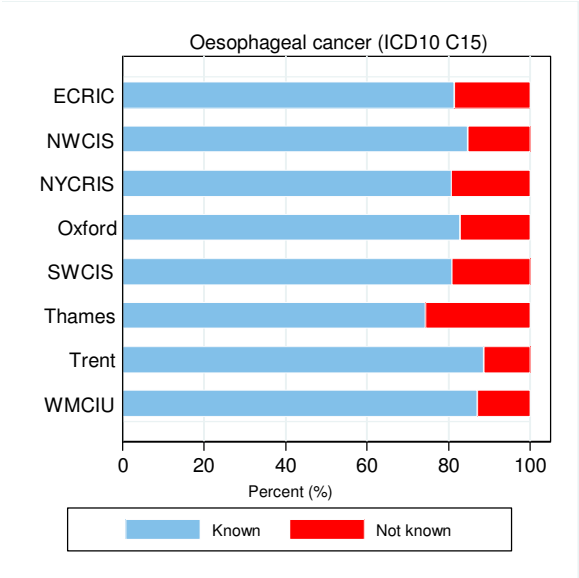
Smaller proportions of cases were microscopically verified in pancreatic cancer (37.4 % in NWCIS to 50.3% in Thames), primary liver cancer (42.4% in NWCIS to 53.3% in Thames), gallbladder cancer (64.1% NWCIS to 76.5% in WMCIU) and biliary cancer (5.92% in ECRIC to 81.4% in Oxford) compared to oesophageal cancer (> 88.9% in all registries), stomach cancer (> 88.7% in all registries) and duodenal cancer (> 85.4% in all registries).

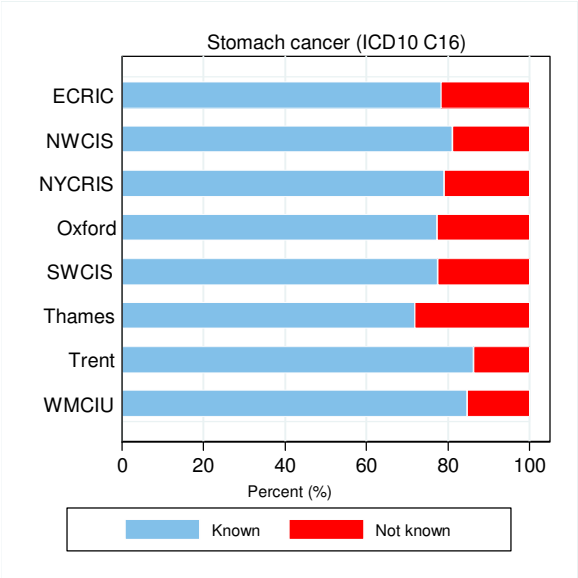
This is not surprising as it is easier to biopsy oesophageal and stomach tumours compared to the more inaccessible hepatic, pancreatic and biliary cancers. Instead primary liver, gallbladder and biliary cancers had a larger proportion of other tests, which included Computed Tomography (CT) scans and X-rays, compared to oesophageal and stomach cancer patients.

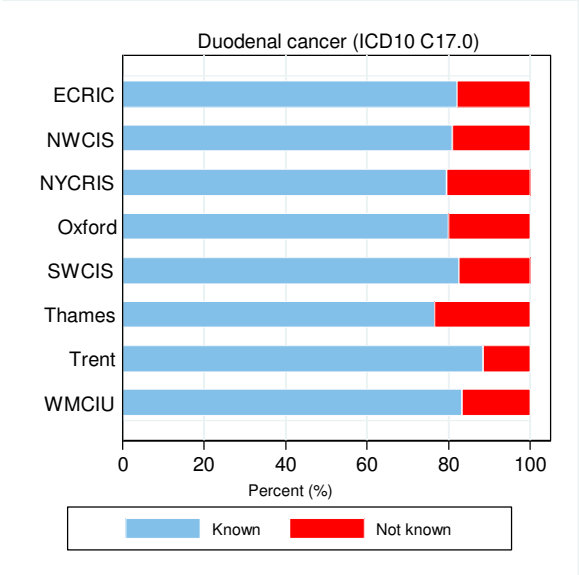
For oesophageal and stomach cancer the highest proportions of missing and unknown basis of diagnoses were in Oxford (3.7% and 5.3% respectively) and Thames (2.8% and 3.1% respectively). These two registries also had the highest proportion of missing and unknown basis of diagnosis for all HPB cancers, particularly for cancers of the pancreas, liver and gallbladder.

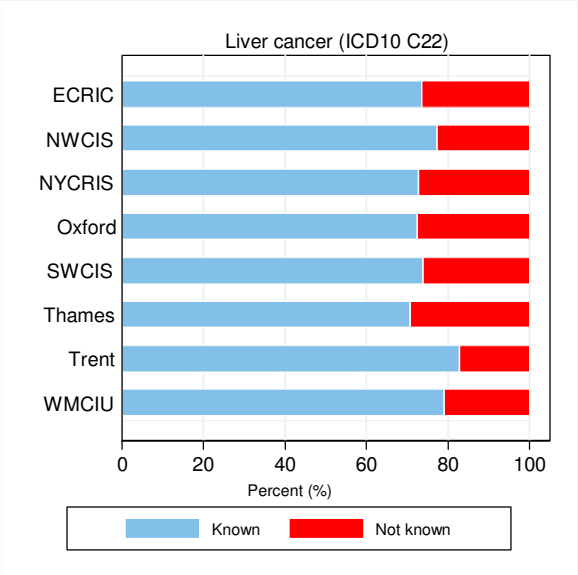
3.5 Proportion of patients with a missing ethnicity by cancer registry

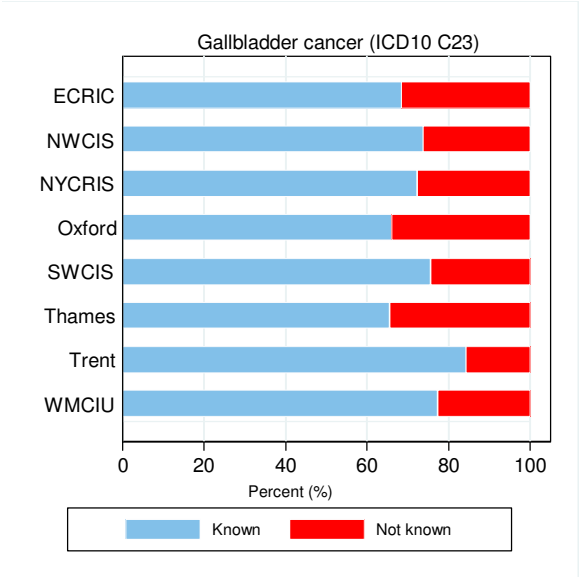
The following graphs show the proportion of patients with missing ethnicity over the period 1998 and 2007. This analysis excludes death certificate only registrations.

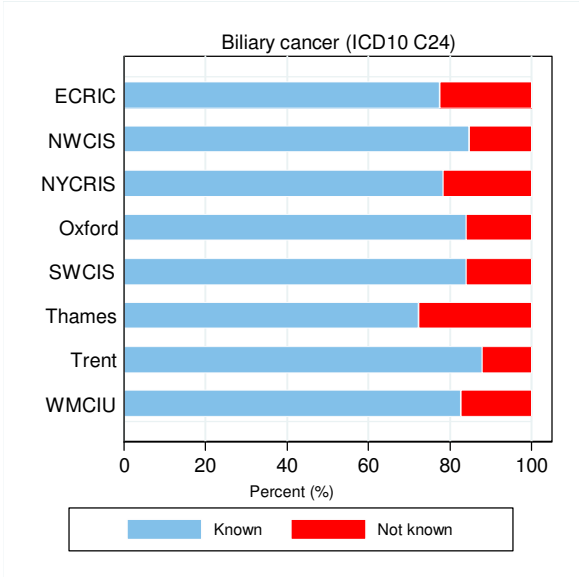


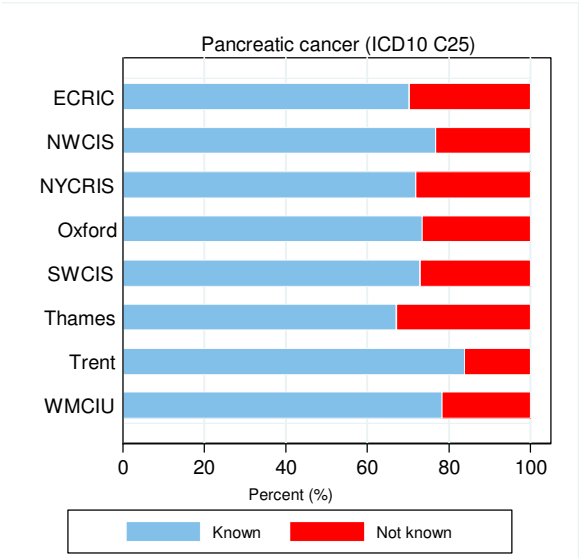












Oesophagus: Missing ethnicity ranged from 11.3% in Trent to 25.6% in Thames.

Stomach: Missing ethnicity ranged from 13.7% in Trent to 28.2% in Thames.

Duodenum: Missing ethnicity ranged from 11.5% in Trent to 23.5% in Thames.

Liver: Missing ethnicity ranged from 17.2% in Trent to 29.3% in Thames.

Gallbladder: Missing ethnicity ranged from 15.8% in Trent to 34.3% in Thames.

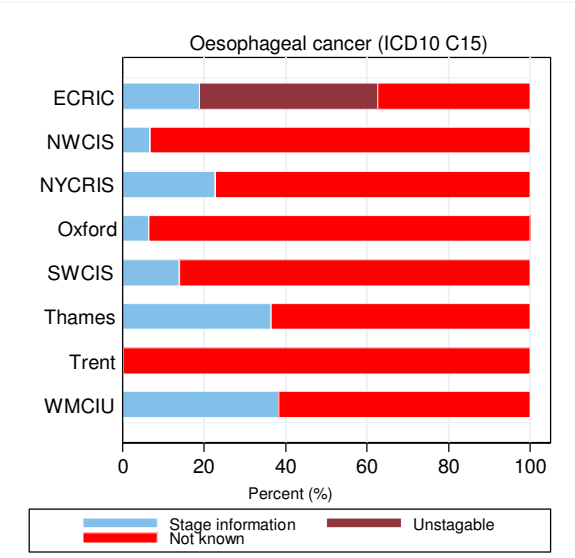
Biliary: Missing ethnicity ranged from 12.2% in Trent to 27.7% in Thames.

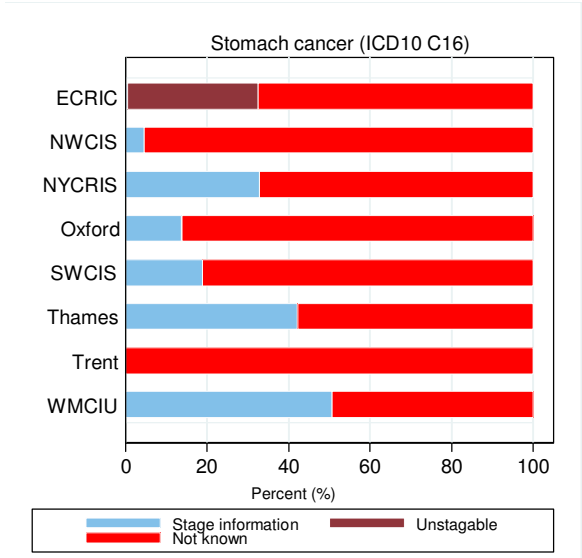
Pancreas: Missing ethnicity ranged from 16.2% in Trent to 32.9% in Thames.

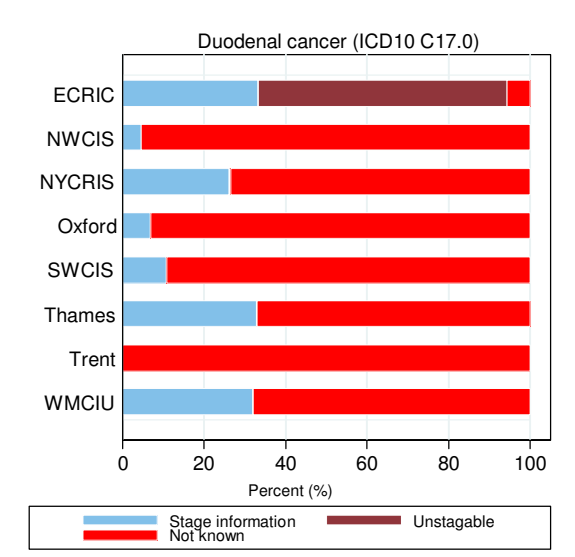
Across all seven cancer groups Trent had the lowest proportion of cases with an unknown ethnicity and Thames had the highest. Less than 21% of patients with oesophageal, stomach, biliary and duodenal cancers had missing ethnicity. Cancers of the gallbladder (26.9%), pancreas (26.6%), and primary liver (25.1%) had the highest proportions of patients with a missing ethnicity.

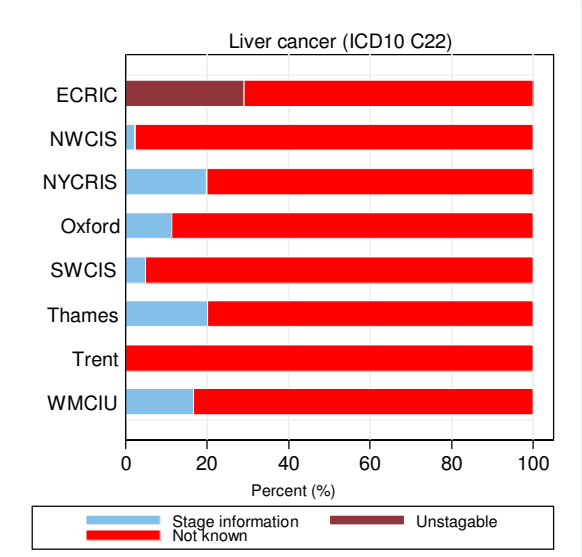
3.6 Proportion of patients with staging information by cancer registry

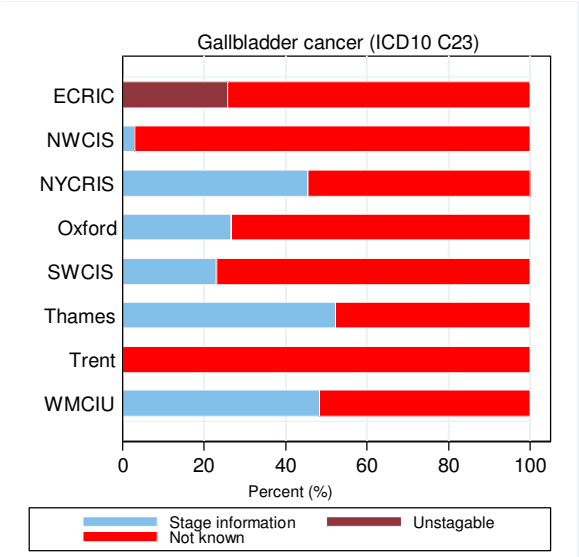
The following graphs show the proportion of patient records with staging information over the period 1998 and 2007. This analysis excludes death certificate only registrations.

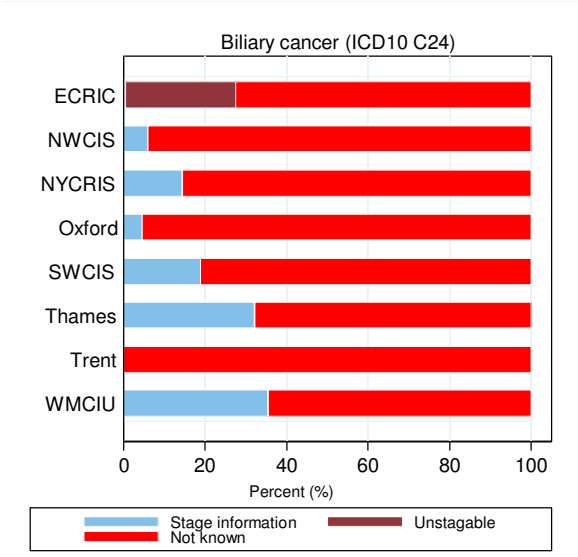


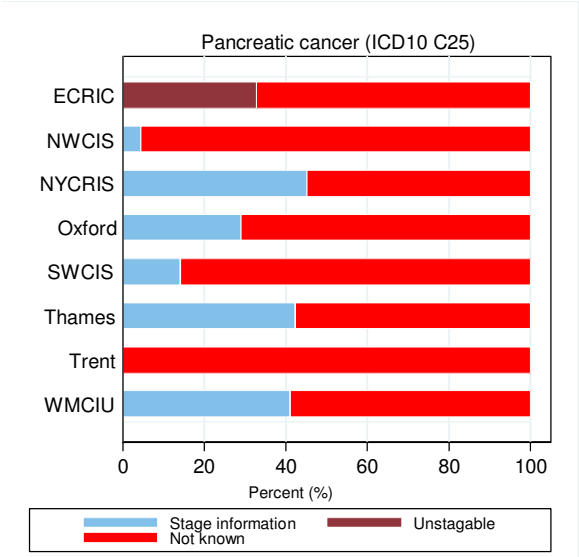












Highest proportion with staging information:

Oesophagus: WMCIU (38.3%) and Thames (36.3%).

Stomach: WMCIU (50.8%), Thames (42.3%) and NYCRIS (32.8%).

Duodenum: Thames (32.8%) and WMCIU (32.0%).

Liver: Thames (20.2%), NYCRIS (20.0%) and WMCIU (16.6%).

Gallbladder: Thames (52.2%), WMCIU (48.4%) and NYCRIS (45.6%).

Biliary: WMCIU (34.5%) and Thames (32.2%).

Pancreas: NYCRIS (45.1%), Thames (42.5%) and WMCIU (41.2%).

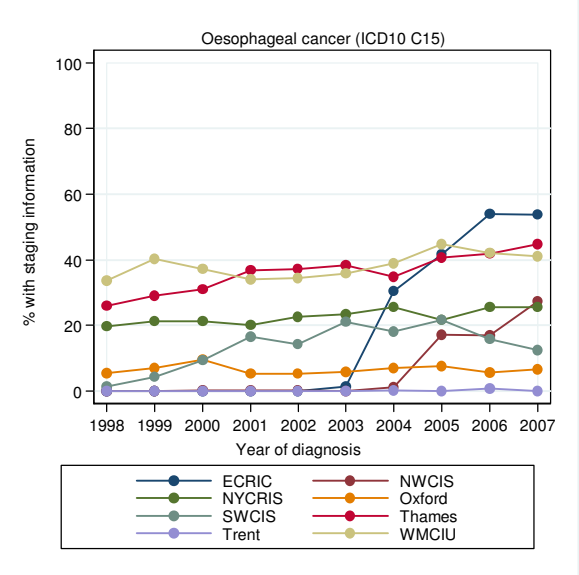
Trent had a very low proportion of staging information for all of the cancer groups in this report.

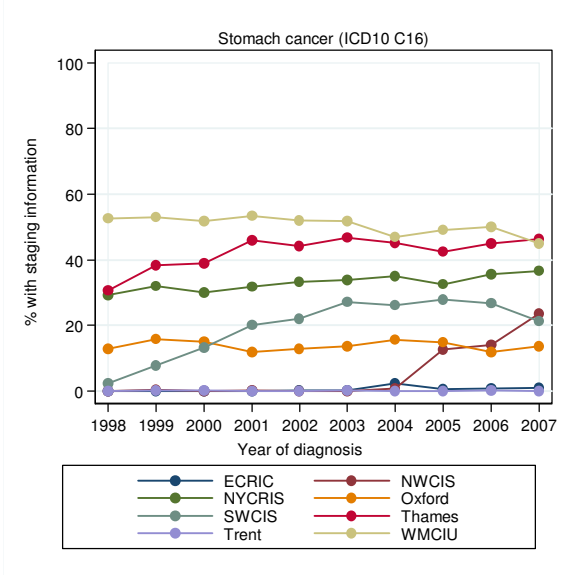
In ECRIC, between 29.0% (liver cancer) and 60.9% (duodenal cancer) of patients could not be staged either due to insufficient information or sufficiently unusual histology at that particular site.

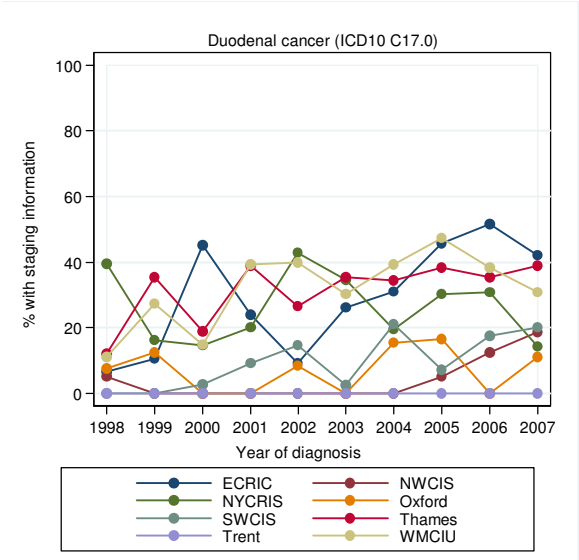
The availability of stage information was poor across all cancer groups. Over three quarters of patients had a missing or unknown stage. Gallbladder cancer had the highest proportion of patients with available stage information (27.2%) and liver cancer the least (10.5%).

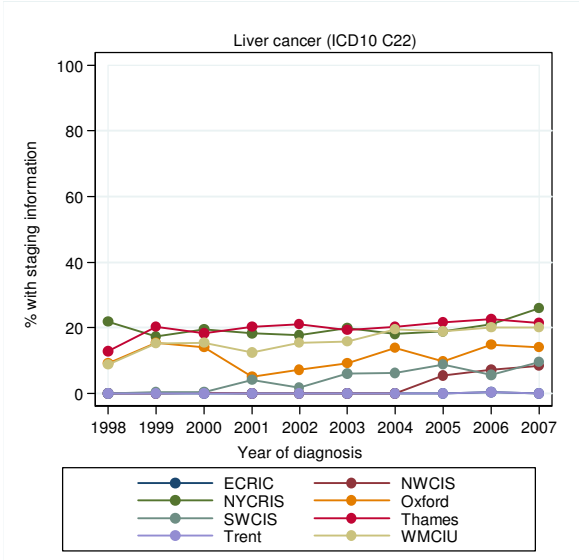
3.7 Proportion of patients with staging information by year and cancer registry

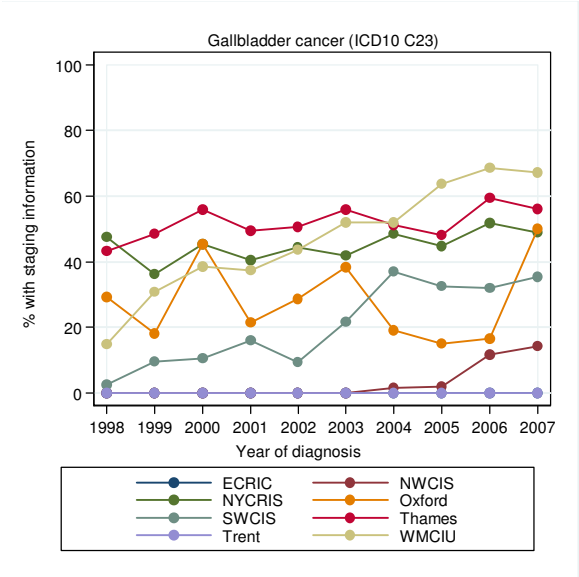
The following graphs show the proportion of patients with staging information by year in each cancer registry. This analysis excludes death certificate only registrations.

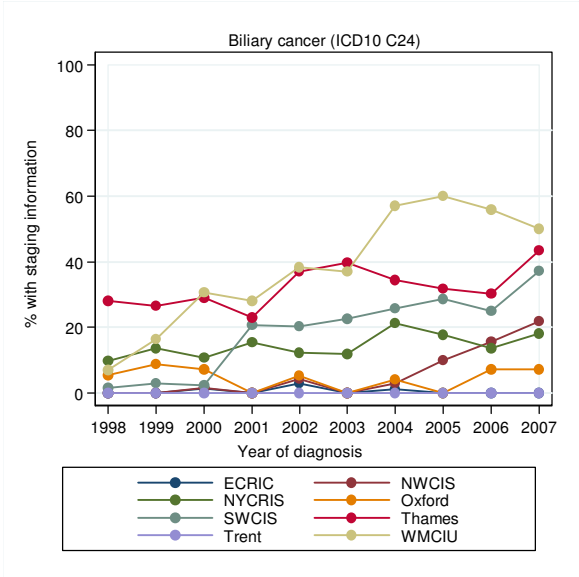


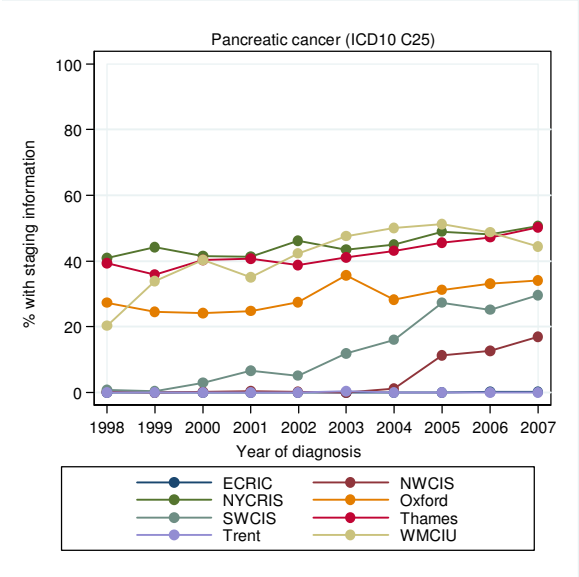












In general, the availability of staging information improved between 1998 and 2006.

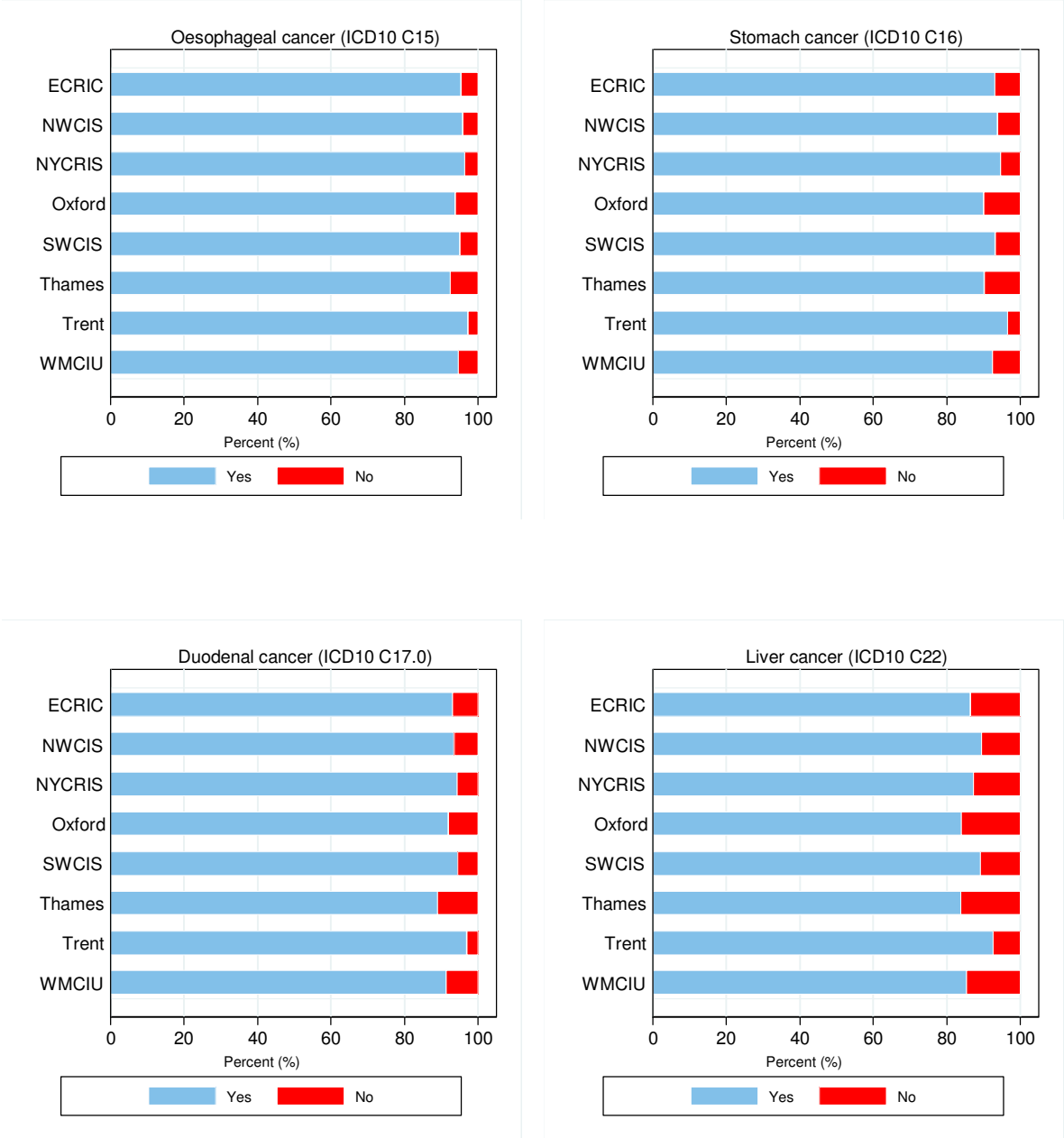
From 2004, the availability of staging information increased in NWCIS across all cancer sites.

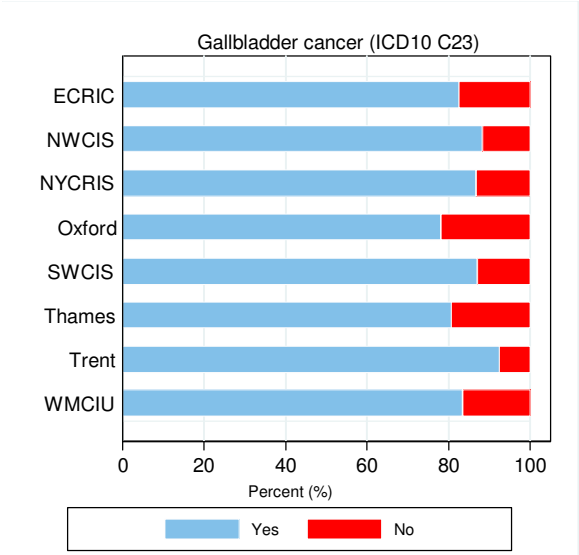
From 2002, the proportion of patients with oesophageal and duodenal cancer with staging information increased in ECRIC.

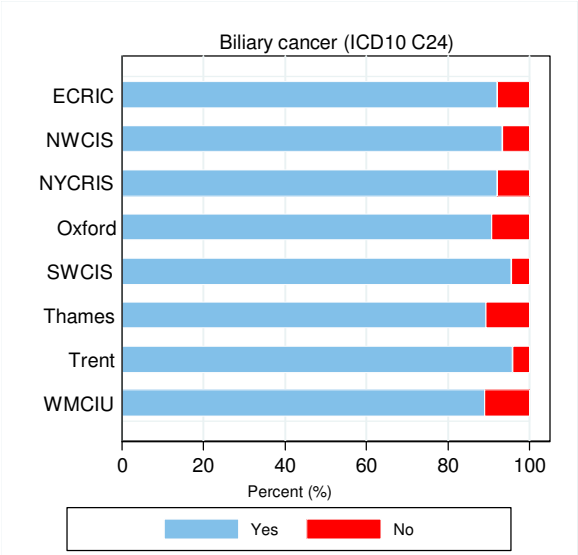
WMCIU and Thames had the highest proportion of available staging information in most cancer groups.

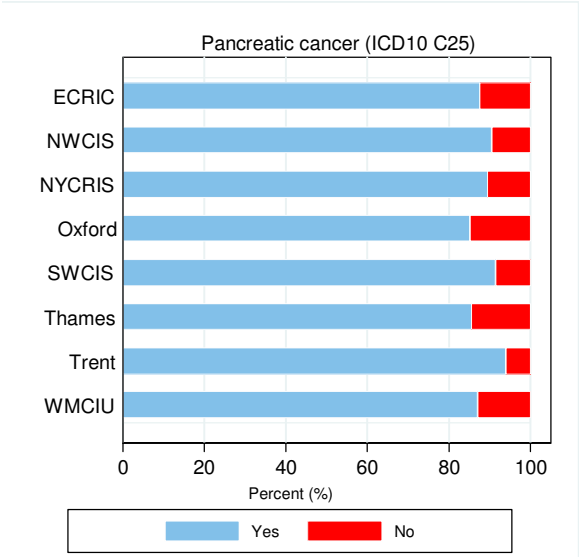
3.8 Proportion of patients with no linked HES record by cancer registry

The following graphs show the proportion of patients with no linked HES record over the period 1998 and 2007. This analysis excludes death certificate only registrations.









Highest proportion with no linked HES records:

Oesophagus: Thames (7.6%) and Oxford (6.2%).

Stomach: Oxford (9.9%), Thames (9.8%) and WMCIU (7.6%).

Duodenum: Thames (11.1%), WMCIU (8.7%) and Oxford (8.0%).

Liver: Thames (16.2%), Oxford (16.0%) and WMCIU (14.6%).

Gallbladder: Oxford (21.9%), Thames (19.3%), ECRIC (17.5%) and WMCIU (16.6%).

Biliary: WMCIU (11.0%) and Thames (10.8%).

Pancreas: Oxford (14.7%), Thames (14.4%) and WMCIU (13.0%).

Gallbladder (14.6%), primary liver (12.7%) and pancreatic cancer (11.1%) had the highest proportion of patients without a linked HES record. Under 10% of the other cancer groups had no linked HES record; oesophagus (5.0%), stomach (6.8%), duodenum (7.0%) and biliary (7.9%).

3.9 Completeness

Table 2 shows the estimated incompleteness of the oesophageal and gastric cancer datasets. In total, only 413 (0.7%) patients with oesophageal cancer and 219 (0.3%) patients with stomach cancer over the ten year period (1998-2007) were estimated to have been potentially missed by the cancer registration process. The lowest completeness appeared to be in the younger age groups and in the earlier years of diagnosis. NWCIS and Thames had the lowest completeness.

Table 2: Completeness of oesophageal and gastric cancer dataset

	Oesophagel cancer (ICD10 C15)			Stomach cancer (ICD10 C16)		
	Cancer registry dataset	HES-onlys	(%)	Cancer registry dataset	HES-onlys	(%)
Total	61,853	413	(0.7)	71,921	219	(0.3)
Males	39,034	292	(0.7)	46,311	139	(0.3)
Females	22,819	120	(0.5)	25,610	78	(0.3)
Missing		1			2	
<50	2,568	43	(1.7)	3,204	25	(0.8)
50-54	3,148	37	(1.2)	2,370	10	(0.4)
55-59	5,058	62	(1.2)	3,996	18	(0.5)
60-64	6,390	67	(1.0)	5,965	25	(0.4)
65-69	7,795	75	(1.0)	9,016	46	(0.5)
70-74	9,564	65	(0.7)	11,910	39	(0.3)
75-79	10,630	52	(0.5)	13,534	42	(0.3)
80-84	8,594	11	(0.1)	11,297	8	(0.1)
85+	8,106	1	(0.0)	10,629	6	(0.1)
1998	5,695	85	(1.5)	8,220	50	(0.6)
1999	5,850	57	(1.0)	7,867	48	(0.6)
2000	6,016	49	(0.8)	7,944	26	(0.3)
2001	6,132	51	(0.8)	7,489	24	(0.3)
2002	6,160	35	(0.6)	7,373	22	(0.3)
2003	6,281	31	(0.5)	6,930	16	(0.2)
2004	6,238	31	(0.5)	6,762	9	(0.1)
2005	6,461	21	(0.3)	6,584	6	(0.1)
2006	6,478	23	(0.4)	6,373	18	(0.3)
2007	6,542	30	(0.5)	6,379		(0.0)
ECRIC	6,257	27	(0.4)	7,188	19	(0.3)
NWCIS	9,250	85	(0.9)	11,125	32	(0.3)
NYCRIS	8,214	36	(0.4)	11,768	33	(0.3)
Oxford	3,009	15	(0.5)	2,763	7	(0.3)
SWCIS	9,567	63	(0.7)	9,429	24	(0.3)
Thames	12,164	92	(0.8)	12,973	59	(0.5)
Trent	6,565	27	(0.4)	8,073	16	(0.2)
WMCIU	6,827	37	(0.5)	8,602	19	(0.2)
Missing		31			10	

Table 3 shows the estimated incompleteness of the hepatic, pancreatic and biliary cancer datasets. In total, only 163 (0.7%) patients with liver cancer, 40 (0.6%) patients with biliary cancer and 387 (0.6%) patients with pancreatic cancer over the ten year period (1998-2007) were estimated to have been potentially missed by the cancer registration process. The lowest completeness appeared to be in the younger age groups. Only 6 / 2,684 (0.2%) patients with duodenal cancer and 4 / 4,550 (0.1%) patients with gallbladder cancer were potentially missed in the cancer registry dataset.

Table 3: Completeness of hepatic, pancreatic and biliary cancer dataset

	Liver cancer (ICD10 C22)			Biliary cancer (ICD10 C24)			Pancreas cancer (ICD10 C25)		
	Cancer registry dataset	HES-onlys	(%)	Cancer registry dataset	HES-onlys	(%)	Cancer registry dataset	HES-onlys	(%)
Total	23,261	163	(0.7)	6,565	40	(0.6)	62,301	387	(0.6)
Males	14,379	83	(0.6)	3,311	25	(0.8)	30,252	211	(0.7)
Females	8,882	79	(0.9)	3,254	15	(0.5)	32,049	175	(0.5)
Missing		1						1	
<50	1,706	70	(4.1)	302	5	(1.7)	2,400	83	(3.5)
50-54	1,163	13	(1.1)	270	4	(1.5)	2,593	34	(1.3)
55-59	1,743	18	(1.0)	479	6	(1.3)	4,354	56	(1.3)
60-64	2,285	19	(0.8)	626	5	(0.8)	6,029	64	(1.1)
65-69	2,975	14	(0.5)	794	10	(1.3)	7,943	56	(0.7)
70-74	3,689	13	(0.4)	1,025	5	(0.5)	9,760	55	(0.6)
75-79	3,945	11	(0.3)	1,097	3	(0.3)	10,819	27	(0.2)
80-84	3,134	1	(0.0)	1,014	1	(0.1)	9,224	11	(0.1)
85+	2,621	4	(0.2)	958	1	(0.1)	9,179	1	(0.0)
1998	1,805	10	(0.6)	665	5	(0.8)	5,671	36	(0.6)
1999	1,882	14	(0.7)	688	2	(0.3)	6,019	30	(0.5)
2000	2,123	9	(0.4)	646	3	(0.5)	6,043	35	(0.6)
2001	2,115	13	(0.6)	623	5	(0.8)	5,985	25	(0.4)
2002	2,293	15	(0.7)	616	5	(0.8)	6,097	39	(0.6)
2003	2,287	13	(0.6)	597	4	(0.7)	6,166	31	(0.5)
2004	2,406	15	(0.6)	590	5	(0.8)	6,456	43	(0.7)
2005	2,660	17	(0.6)	654	3	(0.5)	6,614	37	(0.6)
2006	2,831	32	(1.1)	754	6	(0.8)	6,763	49	(0.7)
2007	2,859	25	(0.9)	732	2	(0.3)	6,487	62	(1.0)
ECRIC	2,003	10	(0.5)	766	5	(0.7)	7,149	34	(0.5)
NWCIS	3,857	15	(0.4)	781	5	(0.6)	7,900	67	(0.8)
NYCRIS	3,338	8	(0.2)	815	5	(0.6)	8,267	44	(0.5)
Oxford	995	6	(0.6)	220	0	(0.0)	3,070	29	(0.9)
SWCIS	3,409	10	(0.3)	1,048	4	(0.4)	9,496	37	(0.4)
Thames	5,027	49	(1.0)	1,505	11	(0.7)	13,524	94	(0.7)
Trent	2,272	7	(0.3)	670	5	(0.7)	6,432	10	(0.2)
WMCIU	2,360	12	(0.5)	760	4	(0.5)	6,463	48	(0.7)
Missing		46			1			24	

3.10 Data quality of oesophago-gastric cancer subgroups

Cancer group	Upper and middle oesophageal cancer			Lower oesophageal cancer			Oesophageal not otherwise specified			Cardia			Distal stomach			Stomach not otherwise specified		
Number of patients	**18,128**			**35,849**			**7,898**			**18,728**			**15,340**			**37,861**		
Death certificate only (DCO)	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**
ECRIC	1,735	1	(0.1)	3,931	10	(0.3)	441	131	(22.9)	2,133	8	(0.4)	1,841	3	(0.2)	2,977	220	(6.9)
NWCIS	2,809	22	(0.8)	4,984	29	(0.6)	1,131	294	(20.6)	2,914	21	(0.7)	1,885	12	(0.6)	5,803	517	(8.2)
NYCRIS	2,576	7	(0.3)	4,844	8	(0.2)	735	107	(12.7)	3,007	11	(0.4)	2,422	10	(0.4)	6,172	183	(2.9)
Oxford	825	0	(0.0)	1,871	0	(0.0)	302	29	(8.8)	648	0	(0.0)	435	0	(0.0)	1,656	39	(2.3)
SWCIS	2,501	7	(0.3)	5,760	6	(0.1)	935	256	(21.5)	2,954	19	(0.6)	2,550	4	(0.2)	3,376	426	(11.2)
Thames	3,392	13	(0.4)	6,705	25	(0.4)	1,510	567	(27.3)	2,789	39	(1.4)	2,099	13	(0.6)	7,181	895	(11.1)
Trent	1,858	10	(0.5)	3,888	12	(0.3)	560	172	(23.5)	2,037	9	(0.4)	2,198	2	(0.1)	3,522	260	(6.9)
WMCIU	2,369	3	(0.1)	3,768	8	(0.2)	515	213	(29.3)	2,129	10	(0.5)	1,859	7	(0.4)	4,282	352	(7.6)
All registries	**18,065**	**63**	**(0.3)**	**35,751**	**98**	**(0.3)**	**6,129**	**1,769**	**(22.4)**	**18,611**	**117**	**(0.6)**	**15,289**	**51**	**(0.3)**	**34,969**	**2,892**	**(7.6)**
Ethnicity	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**
ECRIC	1,409	326	(18.8)	3,272	659	(16.8)	290	151	(34.2)	1,769	364	(17.1)	1,457	384	(20.9)	2,214	763	(25.6)
NWCIS	2,423	386	(13.7)	4,284	700	(14.0)	859	272	(24.0)	2,487	427	(14.7)	1,535	350	(18.6)	4,582	1,221	(21.0)
NYCRIS	2,113	463	(18.0)	2,979	865	(22.5)	484	251	(34.1)	2,499	508	(16.9)	1,974	448	(18.5)	4,706	1,466	(23.8)
Oxford	679	146	(17.7)	1,606	265	(14.2)	197	105	(34.8)	539	109	(16.8)	348	87	(20.0)	1,231	425	(25.7)
SWCIS	2,032	469	(18.8)	4,792	968	(16.8)	612	323	(34.5)	2,462	492	(16.7)	2,082	468	(18.4)	2,346	1,030	(30.5)
Thames	2,563	829	(24.4)	5,109	1,596	(23.8)	960	550	(36.4)	2,074	715	(25.6)	1,543	556	(26.5)	5,054	2,127	(29.6)
Trent	1,644	214	(11.5)	3,517	371	(9.5)	433	127	(22.7)	1,817	220	(10.8)	1,957	241	(11.0)	2,921	601	(17.1)
WMCIU	2,101	268	(11.3)	3,299	469	(12.4)	394	121	(23.5)	1,921	208	(9.8)	1,628	231	(12.4)	3,459	823	(19.2)
All registries	**14,964**	**3,101**	**(17.2)**	**28,858**	**5,893**	**(17.0)**	**4,229**	**1,900**	**(31.0)**	**15,568**	**3,043**	**(16.4)**	**12,524**	**2,765**	**(18.1)**	**26,513**	**8,456**	**(24.2)**
Stage	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**
ECRIC	372	1,363	(78.6)	770	3,161	(80.4)	14	427	(96.8)	28	2,105	(98.7)	5	1,836	(99.7)	0	2,977	(100.0)
NWCIS	196	2,613	(93.0)	397	4,587	(92.0)	14	1,117	(98.8)	179	2,735	(93.9)	79	1,806	(95.8)	225	5,578	(96.1)
NYCRIS	444	2,132	(82.8)	1,199	3,645	(75.2)	212	523	(71.2)	913	2,094	(69.6)	711	1,711	(70.6)	2,180	3,992	(64.7)
Oxford	32	793	(96.1)	135	1,736	(92.8)	30	272	(90.1)	69	579	(89.4)	41	394	(90.6)	269	1,387	(83.8)
SWCIS	277	2,224	(88.9)	965	4,795	(83.2)	44	891	(95.3)	654	2,300	(77.9)	626	1,924	(75.5)	396	2,980	(88.3)
Thames	1,040	2,352	(69.3)	2,835	3,870	(57.7)	348	1,162	(77.0)	1,315	1,474	(52.9)	915	1,184	(56.4)	2,875	4,306	(60.0)
Trent	2	1,856	(99.9)	4	3,884	(99.9)	0	560	(100.0)	0	2,037	(100.0)	2	2,196	(99.9)	2	3,520	(99.9)
WMCIU	831	1,538	(64.9)	1,621	2,147	(57.0)	95	420	(81.6)	1,247	882	(41.4)	1,229	630	(33.9)	1,723	2,559	(59.8)
All registries	**3,194**	**14,871**	**(82.3)**	**7,926**	**27,825**	**(77.8)**	**757**	**5,372**	**(87.6)**	**4,405**	**14,206**	**(76.3)**	**3,608**	**11,681**	**(76.4)**	**7,670**	**27,299**	**(78.1)**
	*ECRIC 744 / 1,363 - unstagable			*ECRIC 1,685 / 3,161 - unstagable			*ECRIC 238 / 427 - unstagable			*ECRIC - 259 / 2,105 - unstagable			*ECRIC - 576 / 1,836 - unstagable			*ECRIC - 1,119 / 2,977 - unstagable		
No linked record in HES	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**
ECRIC	1,664	71	(4.1)	3,802	129	(3.3)	356	85	(19.3)	2,053	80	(3.8)	1,740	101	(5.5)	2,672	305	(10.2)
NWCIS	2,717	92	(3.3)	4,827	157	(3.2)	1,014	117	(10.3)	2,800	114	(3.9)	1,789	96	(5.1)	5,357	446	(7.7)
NYCRIS	2,511	65	(2.5)	4,726	118	(2.4)	615	120	(16.3)	2,933	74	(2.5)	2,331	91	(3.8)	5,718	454	(7.4)
Oxford	778	47	(5.7)	1,786	85	(4.5)	248	54	(17.9)	608	40	(6.2)	412	23	(5.3)	1,448	208	(12.6)
SWCIS	2,408	93	(3.7)	5,546	214	(3.7)	791	144	(15.4)	2,855	99	(3.4)	2,457	93	(3.6)	2,970	406	(12.0)
Thames	3,165	227	(6.7)	6,310	395	(5.9)	1,253	257	(17.0)	2,615	174	(6.2)	1,926	173	(8.2)	6,343	838	(11.7)
Trent	1,821	37	(2.0)	3,815	73	(1.9)	500	60	(10.7)	2,001	36	(1.8)	2,161	37	(1.7)	3,325	197	(5.6)
WMCIU	2,265	104	(4.4)	3,589	179	(4.8)	440	75	(14.6)	2,044	85	(4.0)	1,748	111	(6.0)	3,850	432	(10.1)
All registries	**17,329**	**736**	**(4.1)**	**34,401**	**1,350**	**(3.8)**	**5,217**	**912**	**(14.9)**	**17,909**	**702**	**(3.8)**	**14,564**	**725**	**(4.7)**	**31,683**	**3,286**	**(9.4)**

Note: Codes used to define the OG subgroups are listed in Appendix 1

3.11 Data quality of hepatic, pancreatic and biliary cancer subgroups

Cancer group	Pancreas			Ampulla of Vater			Biliary			Liver			Gallbladder			Duodenum		
Number of patients	**62,310**			**3,258**			**12,638**			**13,939**			**4,550**			**2,684**		
Death certificate only (DCO)	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**	**No**	**Yes**	**(% DCO)**
ECRIC	6,708	438	(6.1)	330	5	(1.5)	1,199	60	(4.8)	1,097	86	(7.3)	457	16	(3.4)	246	4	(1.6)
NWCIS	6,911	1,015	(12.8)	471	17	(3.5)	1,739	205	(10.5)	1,894	348	(15.5)	521	74	(12.4)	369	20	(5.1)
NYCRIS	8,008	265	(3.2)	490	3	(0.6)	1,795	29	(1.6)	1,749	80	(4.4)	654	22	(3.3)	373	2	(0.5)
Oxford	2,952	150	(4.8)	146	0	(0.0)	559	16	(2.8)	499	41	(7.6)	183	3	(1.6)	100	2	(2.0)
SWCIS	8,360	1,074	(11.4)	524	11	(2.1)	1,674	130	(7.2)	1,870	216	(10.4)	542	64	(10.6)	506	14	(2.7)
Thames	12,008	1,552	(11.4)	600	9	(1.5)	2,287	236	(9.4)	2,967	465	(13.5)	810	83	(9.3)	469	23	(4.7)
Trent	5,888	504	(7.9)	362	7	(1.9)	1,272	73	(5.4)	1,062	120	(10.2)	545	33	(5.7)	269	4	(1.5)
WMCIU	5,298	1,179	(18.2)	273	10	(3.5)	1,086	278	(20.4)	1,156	289	(20.0)	459	84	(15.5)	265	18	(6.4)
All registries	**56,133**	**6,177**	**(9.9)**	**3,196**	**62**	**(1.9)**	**11,611**	**1,027**	**(8.1)**	**12,294**	**1,645**	**(11.8)**	**4,171**	**379**	**(8.3)**	**2,597**	**87**	**(3.2)**
Ethnicity	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**	**known**	**Not known**	**(% missing ethnicity)**
ECRIC	4,715	1,993	(29.7)	266	64	(19.4)	910	289	(24.1)	785	312	(28.4)	313	144	(31.5)	202	44	(17.9)
NWCIS	5,299	1,612	(23.3)	403	68	(14.4)	1,405	334	(19.2)	1,425	469	(24.8)	384	137	(26.3)	299	70	(19.0)
NYCRIS	5,756	2,252	(28.1)	405	85	(17.3)	1,327	468	(26.1)	1,251	498	(28.5)	473	181	(27.7)	297	76	(20.4)
Oxford	2,170	782	(26.5)	129	17	(11.6)	424	135	(24.2)	347	152	(30.5)	121	62	(33.9)	80	20	(20.0)
SWCIS	6,100	2,260	(27.0)	456	68	(13.0)	1,301	373	(22.3)	1,353	517	(27.6)	410	132	(24.4)	418	88	(17.4)
Thames	8,059	3,949	(32.9)	479	121	(20.2)	1,617	670	(29.3)	2,068	899	(30.3)	532	278	(34.3)	359	110	(23.5)
Trent	4,932	956	(16.2)	328	34	(9.4)	1,077	195	(15.3)	859	203	(19.1)	459	86	(15.8)	238	31	(11.5)
WMCIU	4,149	1,149	(21.7)	246	27	(9.9)	884	202	(18.6)	883	273	(23.6)	355	104	(22.7)	221	44	(16.6)
All registries	**41,180**	**14,953**	**(26.6)**	**2,712**	**484**	**(15.1)**	**8,945**	**2,666**	**(23.0)**	**8,971**	**3,323**	**(27.0)**	**3,047**	**1,124**	**(26.9)**	**2,114**	**483**	**(18.6)**
Stage	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**	**Stage information**	**Not known**	**(% no stage information)**
ECRIC	3	6,705	(100.0)	4	326	(98.8)	0	1,199	(100.0)	1	1,096	(99.9)	0	457	(100.0)	82	164	(66.7)
NWCIS	312	6,599	(95.5)	24	447	(94.9)	66	1,673	(96.2)	37	1,857	(98.0)	16	505	(96.9)	17	352	(95.4)
NYCRIS	3,618	4,390	(54.8)	50	440	(89.8)	463	1,332	(74.2)	251	1,498	(85.6)	298	356	(54.4)	98	275	(73.7)
Oxford	860	2,092	(70.9)	1	145	(99.3)	73	486	(86.9)	47	452	(90.6)	49	134	(73.2)	7	93	(93.0)
SWCIS	1,181	7,179	(85.9)	127	397	(75.8)	148	1,526	(91.2)	65	1,805	(96.5)	125	417	(76.9)	55	451	(89.1)
Thames	5,100	6,908	(57.5)	204	396	(66.0)	635	1,652	(72.2)	519	2,448	(82.5)	423	387	(47.8)	154	315	(67.2)
Trent	2	5,886	(100.0)	0	362	(100.0)	0	1,272	(100.0)	1	1,061	(99.9)	0	545	(100.0)	0	269	(100.0)
WMCIU	2,182	3,116	(58.8)	131	142	(52.0)	247	839	(77.3)	165	991	(85.7)	222	237	(51.6)	85	180	(67.9)
All registries	**13,258**	**42,875**	**(76.4)**	**541**	**2,655**	**(83.1)**	**1,632**	**9,979**	**(85.9)**	**1,086**	**11,208**	**(91.2)**	**1,133**	**3,038**	**(72.8)**	**498**	**2,099**	**(80.8)**
	*ECRIC 2,202 / 6,705 - unstagable			*ECRIC 78 / 326 - unstagable			*ECRIC 347 / 1,199 - unstagable			*ECRIC - 323 / 1,096 - unstagable			*ECRIC - 118 / 457 - unstagable			*ECRIC - 150 / 164 - unstagable		
No linked record in HES	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**	**Linked**	**Not linked**	**(% no link to HES)**
ECRIC	5,873	835	(12.4)	312	18	(5.5)	1,086	113	(9.4)	914	183	(16.7)	377	80	(17.5)	229	17	(6.9)
NWCIS	6,257	654	(9.5)	446	25	(5.3)	1,606	133	(7.6)	1,649	245	(12.9)	460	61	(11.7)	345	24	(6.5)
NYCRIS	7,161	847	(10.6)	469	21	(4.3)	1,600	195	(10.9)	1,493	256	(14.6)	568	86	(13.1)	352	21	(5.6)
Oxford	2,517	435	(14.7)	138	8	(5.5)	486	73	(13.1)	403	96	(19.2)	143	40	(21.9)	92	8	(8.0)
SWCIS	7,654	706	(8.4)	506	18	(3.4)	1,556	118	(7.0)	1,631	239	(12.8)	472	70	(12.9)	478	28	(5.5)
Thames	10,275	1,733	(14.4)	565	35	(5.8)	1,987	300	(13.1)	2,434	533	(18.0)	654	156	(19.3)	417	52	(11.1)
Trent	5,539	349	(5.9)	351	11	(3.0)	1,208	64	(5.0)	959	103	(9.7)	504	41	(7.5)	261	8	(3.0)
WMCIU	4,611	687	(13.0)	257	16	(5.9)	955	131	(12.1)	960	196	(17.0)	383	76	(16.6)	242	23	(8.7)
All registries	**49,887**	**6,246**	**(11.1)**	**3,044**	**152**	**(4.8)**	**10,484**	**1,127**	**(9.7)**	**10,443**	**1,851**	**(15.1)**	**3,561**	**610**	**(14.6)**	**2,416**	**181**	**(7.0)**

Note: Codes used to define the HPB subgroups are listed in Appendix 1

4. Key findings

- The proportion of death certificate only registrations ranged between 0.5% and 20.8%, although typically remained below 10%. Primary liver, gallbladder and pancreatic cancer had the highest proportions of DCO registrations.
- The proportions of patients with an unspecified anatomical site were highest in patients with oesophageal, stomach and pancreatic cancer.
- The proprtions of microscopically verified cases ranged from 37.4% to 95.1%. Smaller proportions of cases were microscopically verified in pancreatic, primary liver, gallbladder and biliary cancers compared to oesophageal, stomach and duodenal cancer. Pancreatic, primary liver and gallbladder cancer had the highest proportion of patients with an unknown or missing basis of diagnosis.
- Less than 21% of oesophageal, stomach, biliary and duodenal cancers had a missing ethnicity. Cancers of the gallbladder (26.9%), pancreas (26.6%), and primary liver (25.1%) had the highest proportions of patients with a missing ethnicity.
- The availability of stage information was poor across all cancer groups. Over three quarters of patients had a missing or unknown stage. Gallbladder cancer had the highest proportion of patients with available stage information (27.2%) and liver cancer the least (10.5%). In general, the availability of staging information improved between 1998 and 2007.
- Gallbladder (14.6%), primary liver (12.7%) and pancreatic cancer (11.1%) had the highest proportion of patients without a linked HES record. Under 10% of the other cancer groups had no linked HES record; oesophagus (5.0%), stomach (6.8%), duodenum (7.0%) and biliary (7.9%).
- Only small proportions (0.1%-0.7%) of patients with these cancers over the ten year period (1998-2007) were estimated to have been potentially missed by the cancer registration process.

5. Conclusions

This report has investigated the data quality of the records held within NCIN upper gastrointestinal cancer dataset.

The proportion of death certificate only registrations in this dataset was generally low. These registrations would have to be excluded from survival analysis which could potentially bias the results. Although it is unlikely that the number of patients excluded for these cancers would have a major impact on the survival figures, it is important that work continues to reduce the proportion of these registrations.

The proportion of patients with a valid ethnic group classification was high. Only around one fifth to a quarter of patients had missing ethnicity information. With continued improvements in linkage between the NCDR and HES datasets in the future we can hope that the proportion with no ethnicity information will decrease. Also, a high proportion (over 85%) of all cancer groups had a linked record in HES. This will also increase alongside improvements in linkage between the two datasets.

The availability of staging information was poor and this should be improved. However, it is encouraging to note that in general the proportion of patients with staging information is increasing over time. Current work by the UKACR staging sub-group should improve the availability of staging information within the registries. The group's main recommendations include improving engagement with trusts and multi disciplinary teams, using pathological and clinical expertise to allow a stage to be derived where there is only partial staging information and standardising staging practices between registries. Increased focus on the need to improve staging information will help drive this forward.

This report also shows that better classification of oesophageal, stomach and pancreatic tumours is needed to be able to define more specific groups for analyses.

Encouragingly the completeness analysis identified only a very small proportion of potentially missed registrations. This is important as it is likely to have very little impact on analyses carried out on this dataset.

The next data quality report will investigate the quality of these data variables in the next version of the NCIN upper gastrointestinal cancer dataset. It will compare the results to the findings of this report to measure any changes in the quality of the data. In addition, this future report will also calculate the proportions of patients with an unspecified morphology.

Appendix 1: ICD10 codes used in this report.

Oesophago-gastric group	ICD10 code
Oesophageal cancer	C15
Stomach cancer	C16

Hepatic, pancreatic and biliary group	ICD10 code
Duodendum	C17.0
Liver	C22
Gallbladder	C23
Biliary	C24
Pancreas	C25

More detailed groups

OG groups	ICD10 and morphology codes
Upper / Middle oesophagus	C15.0, C15.1, C15.3, C15.4 including C15.8 & C15.9 with a morphology code of 8050-8083 (Squamous cell carcinomas)
Lower oesophagus	C15.2, C15.5 including C15.8 & C15.9 with a morphology code of 8140-8576 (Adenocarcinomas)
Oesophagus unknown	C15.8, C15.9
Cardia	C16.0
Stomach	C16.1, C16.2, C16.3, C16.4, C16.5, C16.6
Stomach unknown	C16.8, C16.9

HPB groups	ICD10 codes
Duodendum	C17.0
Liver (excluding intrahepatic bile duct)	C22.0, C22.2, C22.3, C22.4, C22.7, C22.9
Intrahepatic bile duct, Bile duct, Biliary tract (cholangiocarcinomas)	C22.1, C24.0, C24.8, C24.9
Gallbladder	C23
Ampulla of Vater	C24.1
Pancreas	C25 (all)

Appendix 2: List of ICD10 4 digit codes

C15 Malignant neoplasm of oesophagus

C15.0 Malignant neoplasm: Cervical part of oesophagus
C15.1 Malignant neoplasm: Thoracic part of oesophagus
C15.2 Malignant neoplasm: Abdominal part of oesophagus
C15.3 Malignant neoplasm: Upper third of oesophagus
C15.4 Malignant neoplasm: Middle third of oesophagus
C15.5 Malignant neoplasm: Lower third of oesophagus
C15.8 Malignant neoplasm: Overlapping lesion of oesophagus
C15.9 Malignant neoplasm: Oesophagus, unspecified

C16 Malignant neoplasm of stomach

C16.0 Malignant neoplasm: Cardia
C16.1 Malignant neoplasm: Fundus of stomach
C16.2 Malignant neoplasm: Body of stomach
C16.3 Malignant neoplasm: Pyloric antrum
C16.4 Malignant neoplasm: Pylorus
C16.5 Malignant neoplasm: Lesser curvature of stomach, unspecified
C16.6 Malignant neoplasm: Greater curvature of stomach, unspecified
C16.8 Malignant neoplasm: Overlapping lesion of stomach
C16.9 Malignant neoplasm: Stomach, unspecified

C17 Malignant neoplasm of small intestine

C17.0 Malignant neoplasm: Duodenum
C17.1 Malignant neoplasm: Jejunum
C17.2 Malignant neoplasm: Ileum
C17.3 Malignant neoplasm: Meckel's diverticulum
C17.8 Malignant neoplasm: Overlapping lesion of small intestine
C17.9 Malignant neoplasm: Small intestine, unspecified

(Not included in the upper gastrointestinal cancer dataset)

C22 Malignant neoplasm of liver and intrahepatic bile ducts

C22.0 Malignant neoplasm: Liver cell carcinoma
C22.1 Malignant neoplasm: Intrahepatic bile duct carcinoma
C22.2 Malignant neoplasm: Hepatoblastoma
C22.3 Malignant neoplasm: Angiosarcoma of liver
C22.4 Malignant neoplasm: Other sarcomas of liver
C22.7 Malignant neoplasm: Other specified carcinomas of liver
C22.9 Malignant neoplasm: Liver, unspecified

C23 Malignant neoplasm of gallbladder

C24 Malignant neoplasm of other and unspecified parts of biliary tract

C24.0 Malignant neoplasm: Extrahepatic bile duct
C24.1 Malignant neoplasm: Ampulla of Vater
C24.8 Malignant neoplasm: Overlapping lesion of biliary tract
C24.9 Malignant neoplasm: Biliary tract, unspecified

C25 Malignant neoplasm of pancreas

C25.0 Malignant neoplasm: Head of pancreas
C25.1 Malignant neoplasm: Body of pancreas
C25.2 Malignant neoplasm: Tail of pancreas
C25.3 Malignant neoplasm: Pancreatic duct
C25.4 Malignant neoplasm: Endocrine pancreas
C25.7 Malignant neoplasm: Other parts of pancreas
C25.8 Malignant neoplasm: Overlapping lesion of pancreas
C25.9 Malignant neoplasm: Pancreas, unspecified

Source: http://apps.who.int/classifications/apps/icd/icd10online/

Appendix 3: List of ICD10 codes and procedure codes used in the completeness analysis.

Oesophageal cancer (ICD10 C15)	G011	Oesophagogastrectomy and anastomosis of oesophagus to stomach
	G018	Other specified excision of oesophagus and stomach
	G019	Unspecified excision of oesophagus and stomach
	G038	Other specified partial excision of oesophagus
	G039	Unspecified partial excision of oesophagus
	G021	Total oesophagectomy and anastomosis of pharynx to stomach
	G022	Total oesophagectomy and interposition of microvascularly attached jejunum
	G023	Total oesophagectomy and interposition of jejunum NEC
	G024	Total oesophagectomy and interposition of microvascularly attached colon
	G025	Total oesophagectomy and interposition of colon NEC
	G031	Partial oesophagectomy and end to end anastomosis of oesophagus
	G032	Partial oesophagectomy and interposition of microvascularly attached jejunum
	G035	Partial oesophagectomy and interposition of microvascularly attached colon
	G036	Partial oesophagectomy and interposition of colon NEC
	G028	Other specified total excision of oesophagus
	G029	Unspecified total excision of oesophagus
Stomach (ICD10 C16)	G012	Oesophagogastrectomy and anastomosis of oesophagus to transposed jejunum
	G013	Oesophagogastrectomy and anastomosis of oesophagus to jejunum NEC
	G271	Total gastrectomy and excision of surrounding tissue
	G272	Total gastrectomy and anastomosis of oesophagus to duodenum
	G273	Total gastrectomy and interposition of jejunum
	G274	Total gastrectomy and anastomosis of oesophagus to transposed jejunum
	G275	Total gastrectomy and anastomosis of oesophagus to jejunum NEC
	G278	Other specified total excision of stomach
Duodenum (ICD10 C17.0)	G491	Gastroduodenectomy
	G492	Total excision of duodenum
	G493	Partial excision of duodenum
	G498	Other specified excision of duodenum
	G499	Unspecified excision of duodenum
Liver (ICD10 C22)	J021	Right hemihepatectomy NEC
	J022	Left hemihepatectomy NEC
	J023	Resection of segment of liver
	J024	Wedge excision of liver
	J026	Extended right hemihepatectomy
	J027	Extended left hemihepatectomy
	J028	Other specified partial excision of liver
	J029	Unspecified partial excision of liver
Gallbladder (ICD10 C23)	J188	Other specified excision of gall bladder
	J189	Unspecified excision of gall bladder
Biliary (ICD10 C24)	J181	Total cholecystectomy and excision of surrounding tissue
	J182	Total cholecystectomy and exploration of common bile duct
	J183	Total cholecystectomy NEC
	J184	Partial cholecystectomy and exploration of common bile duct
	J185	Partial cholecystectomy NEC
	J271	Excision of ampulla of Vater and replantation of common bile duct into duodenum
	J272	Partial excision of bile duct and anastomosis of bile duct to duodenum
	J273	Partial excision of bile duct and anastomosis of bile duct to jejunum
	J274	Partial excision of bile duct and end to end anastomosis of bile duct
	J275	Excision of extrahepatic bile ducts HFQ
	J278	Other specified excision of bile duct
	J279	Unspecified excision of bile duct
Pancreas (ICD10 C25)	J551	Total pancreatectomy and excision of surrounding tissue
	J552	Total pancreatectomy NEC
	J553	Excision of transplanted pancreas
	J558	Other specified total excision of pancreas
	J559	Unspecified total excision of pancreas
	J561	Pancreaticoduodenectomy and excision of surrounding tissue
	J562	Pancreaticoduodenectomy and resection of antrum of stomach
	J563	Pancreaticoduodenectomy NEC
	J564	Subtotal excision of head of pancreas with preservation of duodenum and drainage HFQ
	J568	Other specified excision of head of pancreas
	J569	Unspecified excision of head of pancreas
	J571	Subtotal pancreatectomy
	J572	Left pancreatectomy and drainage of pancreatic duct
	J573	Left pancreatectomy NEC
	J574	Excision of tail of pancreas and drainage of pancreatic duct
	J575	Excision of tail of pancreas NEC
	J578	Other specified other partial excision of pancreas
	J579	Unspecified other partial excision of pancreas

FIND OUT MORE:

Thames Cancer Registry is the lead Cancer Registry for upper gastrointestinal cancers.

The NCIN is a UK-wide initiative, working closely with cancer services in England, Scotland, Wales and Northern Ireland, and the NCRI, to drive improvements in standards of cancer care and clinical outcomes by improving and using the information it collects for analysis, publication and research. In England, the NCIN is part of the National Cancer Programme.