Data quality and completeness report:

Upper Gastrointestinal Site Specific Clinical Reference Group (SSCRG)

Victoria H Coupland
Julie Konfortion
Ruth H Jack
Karen M Linklater

Contents

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 cancers) and primary hepatic, pancreatic and biliary cancers (including cancers of the liver, biliary tract, ampulla of Vater, duodenum, gallbladder and pancreas), (Appendix 1). Thames Cancer Registry investigates these cancers using data from the National Cancer Data Repository (NCDR). The NCDR contains information from the eight English cancer registries on all patients diagnosed with cancer in their respective catchment areas.

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 subgroups would be difficult. It is vital to record the quality of these data to ensure improvements can be made. 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 1999 and 2008 focusing on the most recent diagnosis year (2008).

2. Methods

Data were extracted from the National Cancer Data Repository on all patients diagnosed with primary upper gastrointestinal cancers between 1999 and 2008. There were 146,428 OG tumours and 110,815 HPB tumours diagnosed in this ten-year period. Of these, 12,799 OG tumours and 11,490 HPB tumours were diagnosed in 2008.

2.1 Data quality

The quality of the dataset was investigated for the main cancer types including cancers of the oesophagus (International Classification of Diseases version 10 (ICD10) C15), stomach (ICD10 C16), duodenum (ICD10 C17.0), primary liver (ICD10 C22), gallbladder (ICD10 C23), biliary tract (ICD10 C24) and pancreas (ICD10 C25).

Data were displayed for registrations diagnosed in 2008 by type of cancer, and the trends over time (1999-2008) by type of cancer were also plotted. Finally, data were also analysed at cancer registry level for each cancer type. The graphs and accompanying text will refer to each registry by their code (Table 1).

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) 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 Hospital Episode Statistics (HES) dataset. Many cases are found and their details are updated to form a complete registration. However, some cases may not have been seen in a hospital and therefore further details cannot be found. These will 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) Basis of diagnosis

The basis of diagnosis is recorded for each cancer registration. Five groups were defined as follows: microscopically verified (cytology, histology of primary tumour and histology of metastases), clinically verified (clinical opinion, clinical investigation and specific tumour markers), death certificate, not known and missing.

c) Anatomical site

The unknown anatomical site group included patients with an ICD10 four digit code of Cxx.8 (overlapping lesion of the cancer in question) and Cxx.9 (unspecified anatomical subsite of the cancer in question). See Appendix 1 for a full list of codes. Large proportions of patients with an unspecified anatomical site will limit our ability to analyse these cancers by specific subgroups.

d) Morphology

Large proportions of patients with an unknown morphology code will limit our ability to analyse these cancers by specific morphology subgroups. Morphology was classified as known (valid morphology codes) and not known (morphology codes: 8000, 8001 and missing).

e) Linked HES records

If a registration has no linked HES record this could indicate that the matching was not 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 no corresponding cancer diagnosis coded 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 hospital activity. This will be important to consider in any future treatment analysis.

f) 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 NCDR includes ethnicity from the HES dataset. Large proportions of patients with a missing ethnicity code will make studies focussing on ethnicity less robust.

g) Stage variables

Stage is an important indicator of the prognosis and will influence the treatment that patients receive. The NCDR records TNM staging information. T describes the size of the tumour, N whether regional lymph nodes are involved and M describes distant metastasis. There are three types of TNM staging in the NCDR: pathological TNM (t_path, n_path, m_path, tnm_path), clinical TNM (t_clin, n_clin, m_clin, tnm_clin) and integrated TNM (t_int, n_int, m_int, tnm_int). The NCDR also includes the field “mets” which records if a patient has distant metastases or not and the field “nodes_postive” which records the number of nodes that were found to be positive. Each of these variables were analysed separately, with the proportion of registrations with a valid known or missing code calculated. For the individual T, N, M and “mets” fields a value of X was recorded as valid not known. In the “nodes_positive” field a value of 99 or 999 was defined as valid not known.

2.2 Completeness

The completeness of case ascertainment in the cancer registry has often been questioned. It is important to ascertain an estimate of how many 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 in 2008 and who had no matching record in the cancer registry dataset were identified (HES-onlys). HES-only registrations were then narrowed down to include only those with a relevant surgical procedure code related to the cancer in question (see Appendix 2). 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.

3. Results

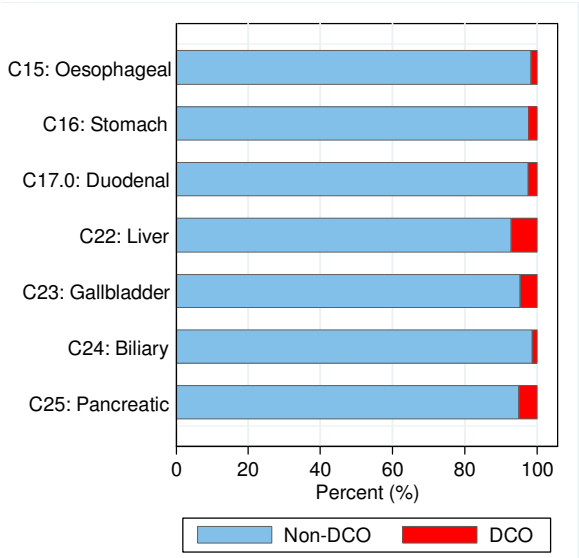
3.1 Quality of the upper gastrointestinal cancer dataset, England, 2008

	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 C25)	
	6,671		6,128		367		2,978		551		777		6,817	
Death certificate only (DCO)														
DCO	104	(3.1)	147	(4.3)	9	(2.5)	209	(7.0)	25	(4.5)	10	(1.3)	344	(5.0)
Non-DCO	6,567	(96.9)	5,981	(95.7)	358	(97.5)	2,769	(93.0)	526	(95.5)	767	(98.7)	6,473	(95.0)
Basis of diagnosis (excluding DCO registrations)														
Microscopically verified	6,115	(93.1)	5,468	(91.4)	300	(83.8)	1,287	(46.5)	345	(65.6)	593	(77.3)	3,100	(47.9)
Clinically verified	419	(6.4)	481	(8.0)	56	(15.6)	1,441	(52.0)	173	(32.9)	167	(21.8)	3,306	(51.1)
Death certificate	9	(0.1)	9	(0.2)	0	(0.0)	16	(0.6)	2	(0.4)	2	(0.3)	17	(0.3)
Not known	13	(0.2)	5	(0.1)	0	(0.0)	9	(0.3)	3	(0.6)	3	(0.4)	23	(0.4)
Missing	11	(0.2)	18	(0.3)	2	(0.6)	16	(0.6)	3	(0.6)	2	(0.3)	27	(0.4)
Anatomical site (excluding DCO registrations)														
Known	3,649	(55.6)	3,077	(51.4)	-	-	-	-	-	-	665	(86.7)	3,032	(46.8)
Not known	2,918	(44.4)	2,904	(48.6)	-	-	-	-	-	-	102	(13.3)	3,441	(53.2)
Morphology (excluding DCO registrations)														
Known	6,399	(97.4)	5,774	(96.5)	336	(93.9)	2,510	(90.6)	456	(86.7)	716	(93.4)	5,327	(82.3)
Not known	168	(2.6)	207	(3.5)	22	(6.1)	259	(9.4)	70	(13.3)	51	(6.6)	1,146	(17.7)
Linked HES records (excluding DCO registrations)														
Linked	6,397	(97.4)	5,735	(95.9)	342	(95.5)	2,530	(91.4)	468	(89.0)	734	(95.7)	5,941	(91.8)
Not linked	170	(2.6)	246	(4.1)	16	(4.5)	239	(8.6)	58	(11.0)	33	(4.3)	532	(8.2)
Ethnicity (excluding DCO registrations)														
Known	6,138	(93.5)	5,511	(92.1)	326	(91.1)	2,421	(87.4)	446	(84.8)	711	(92.7)	5,625	(86.9)
Not known	429	(6.5)	470	(7.9)	32	(8.9)	348	(12.6)	80	(15.2)	56	(7.3)	848	(13.1)

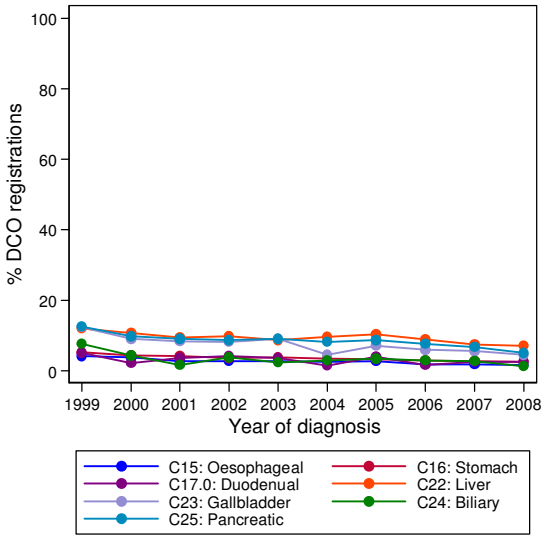
3.2 Death certificate only (DCO)

The following graphs show the proportion of death certificate only registrations for each cancer type.

2008

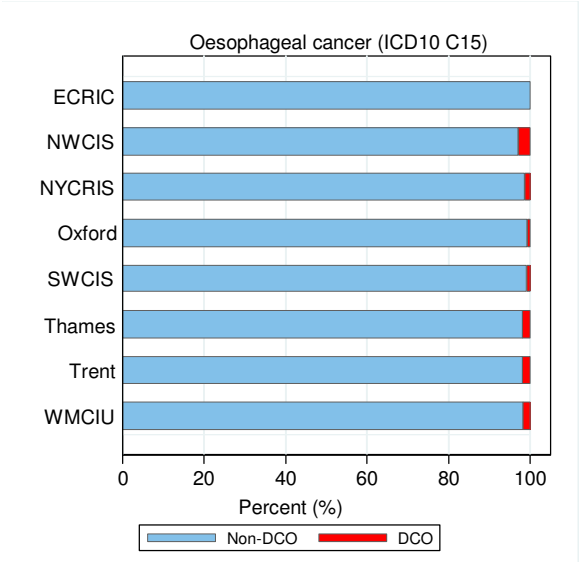


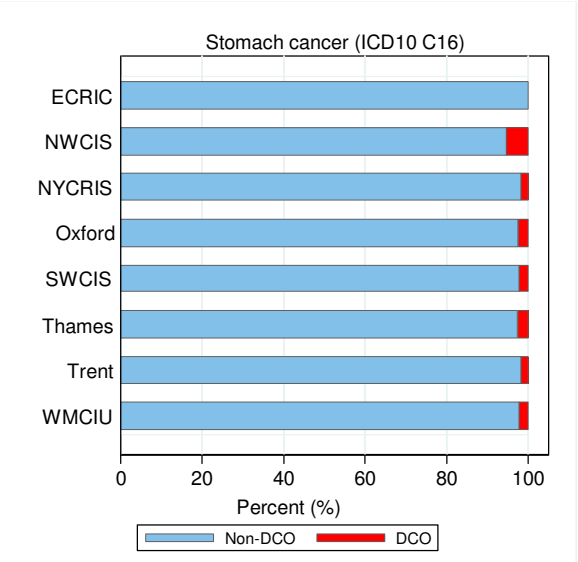
1999-2008

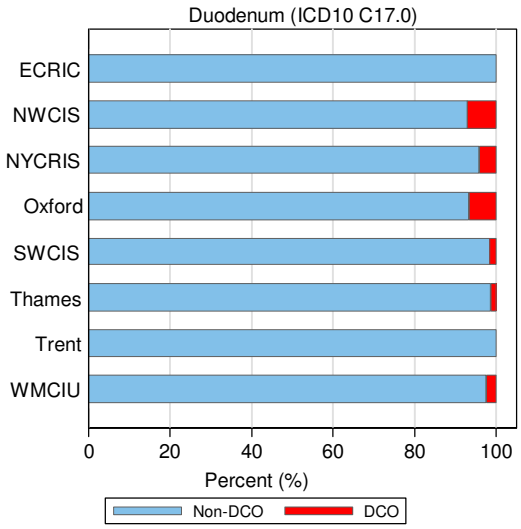


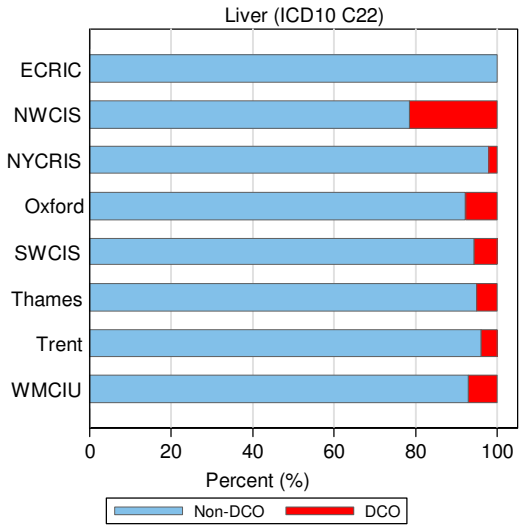
Less than 5% of oesophageal, stomach, duodenal, gallbladder and biliary cancer registrations were based on the death certificate only. The greatest proportion of DCO registrations was in liver cancer (7%). Between 1999 and 2008, the proportion of DCO registrations decreased for all cancer types.

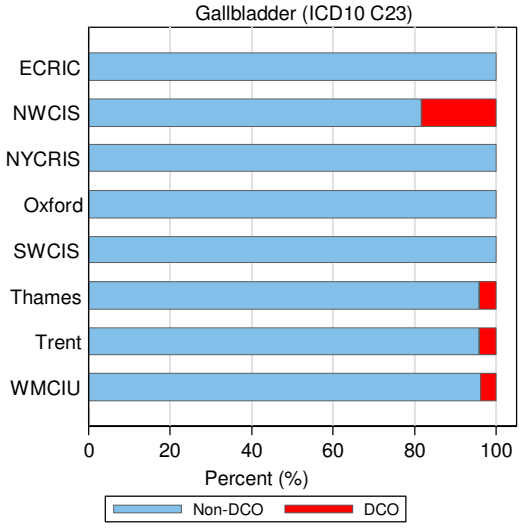
2008

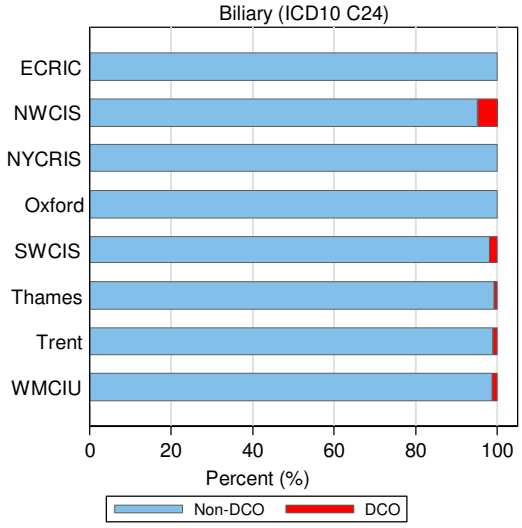


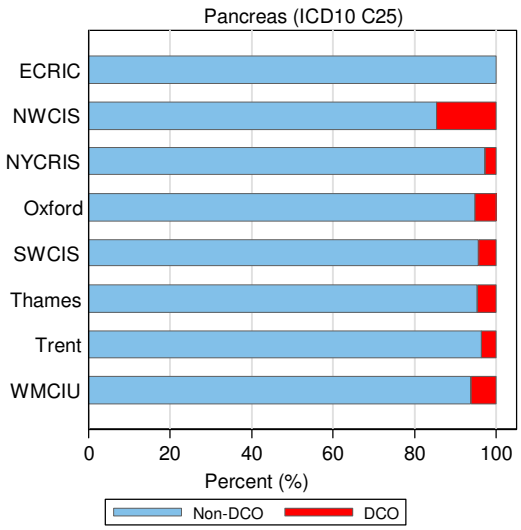










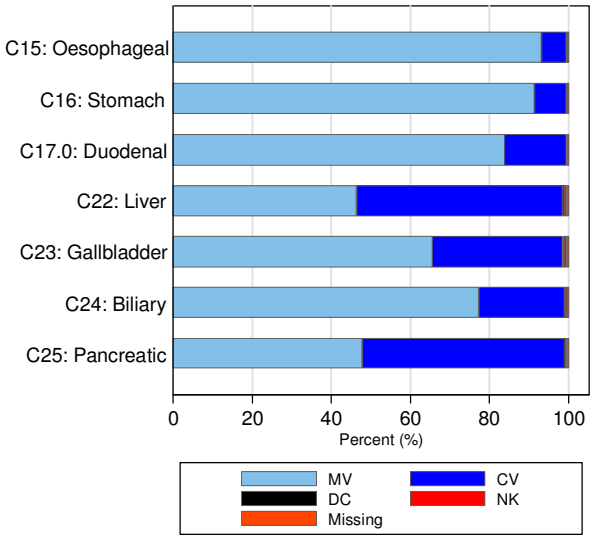


For most of the cancer types the proportion of DCO registrations was very low, with little variation between cancer registries.

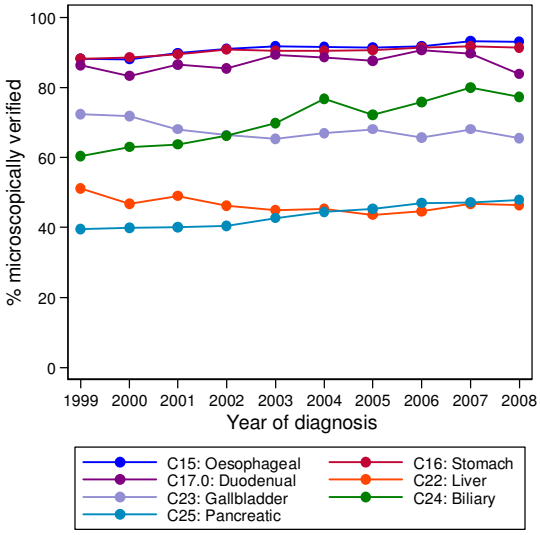
3.3 Basis of diagnosis

The following graphs show the proportion of registrations whose basis of diagnosis was either microscopically verified (MV), clinically verified (CV), death certificate (DC), not known (NK) or missing for each cancer type. This analysis excludes death certificate only registrations.

2008

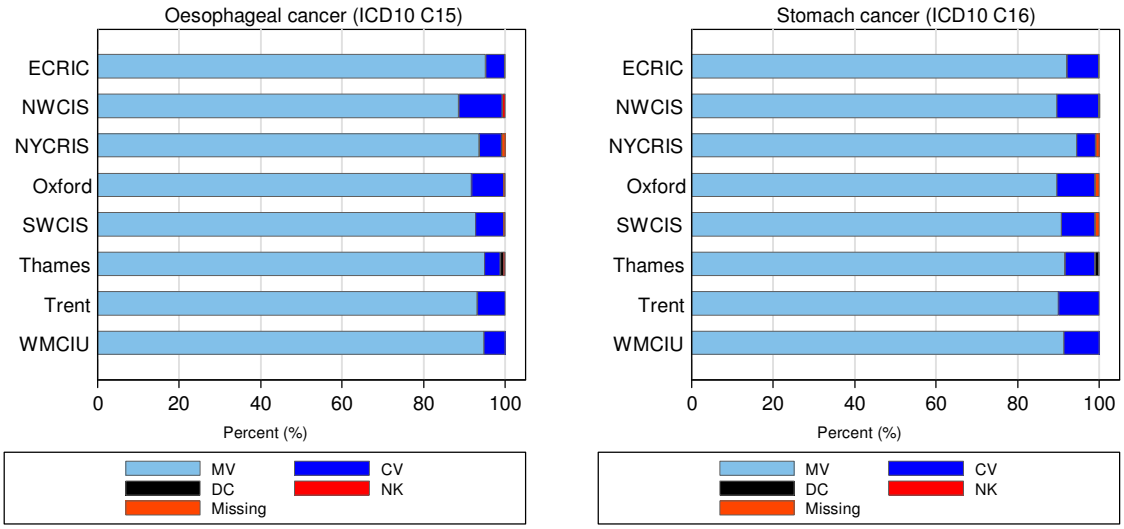


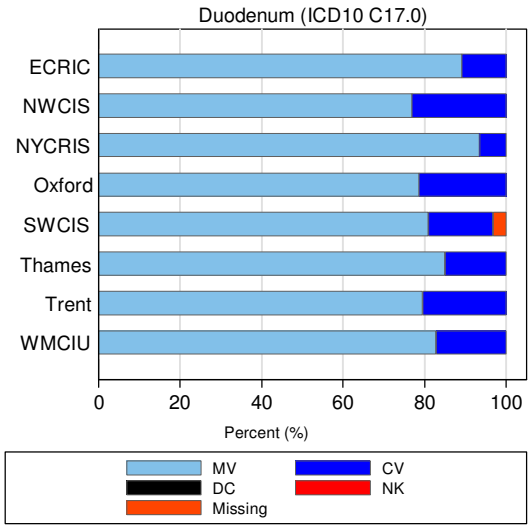
1999-2008

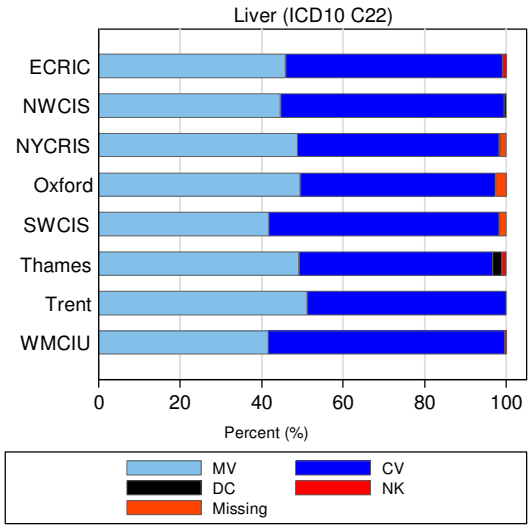


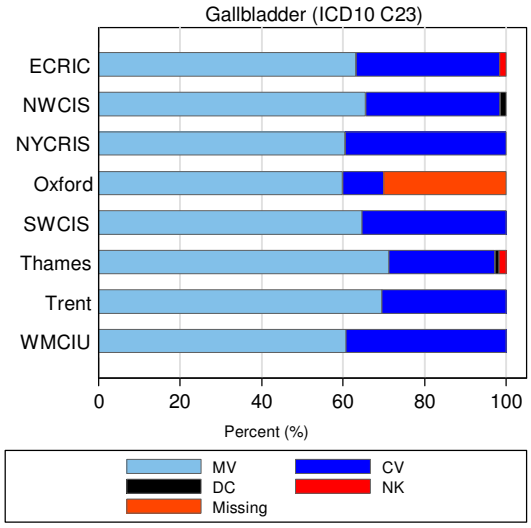
Over 90% of oesophageal and stomach cancers were microscopically verified. Liver (47%) and pancreatic (48%) cancers had the lowest proportions of microscopically verified cases. Around 50% of these cancers were clinically verified. Between 1999 and 2008 there was an increase in the proportion of biliary and pancreatic cancers and a decrease in the proportion of gallbladder and liver cancers that were microscopically verified.

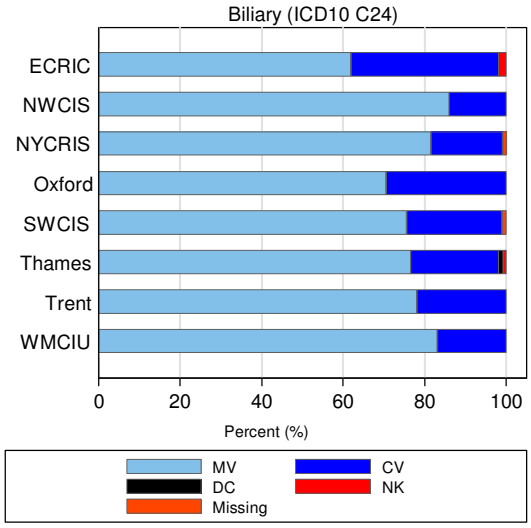
2008

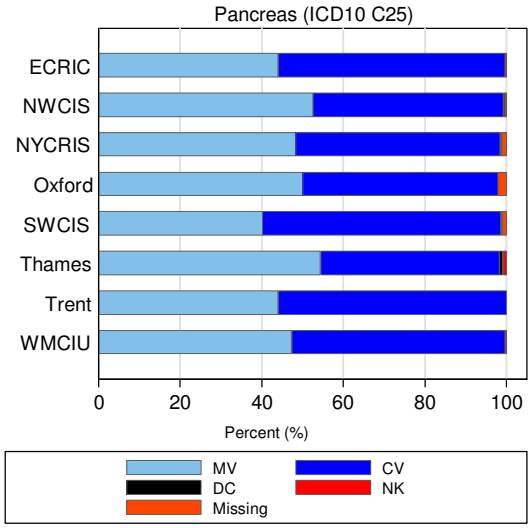










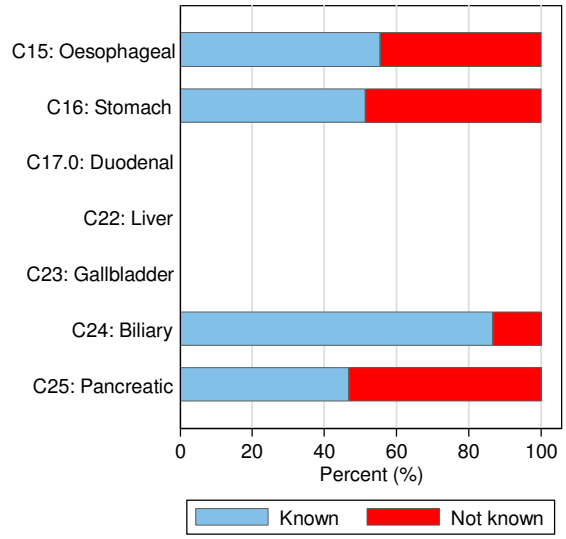


The lower proportion of microscopically verified hepatic, pancreatic and biliary cancers is probably due to these tumours being more inaccessible compared with the more accessible oesophageal and stomach cancers.

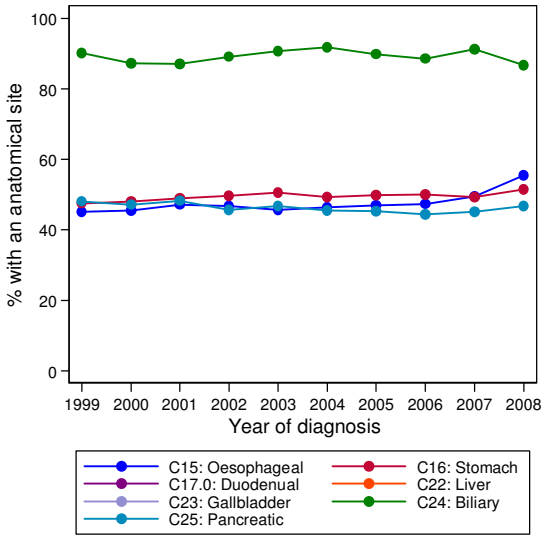
3.4 Anatomical site

The following graphs show the proportion of registrations with known and not known anatomical subsites. This analysis excludes death certificate only registrations.

2008

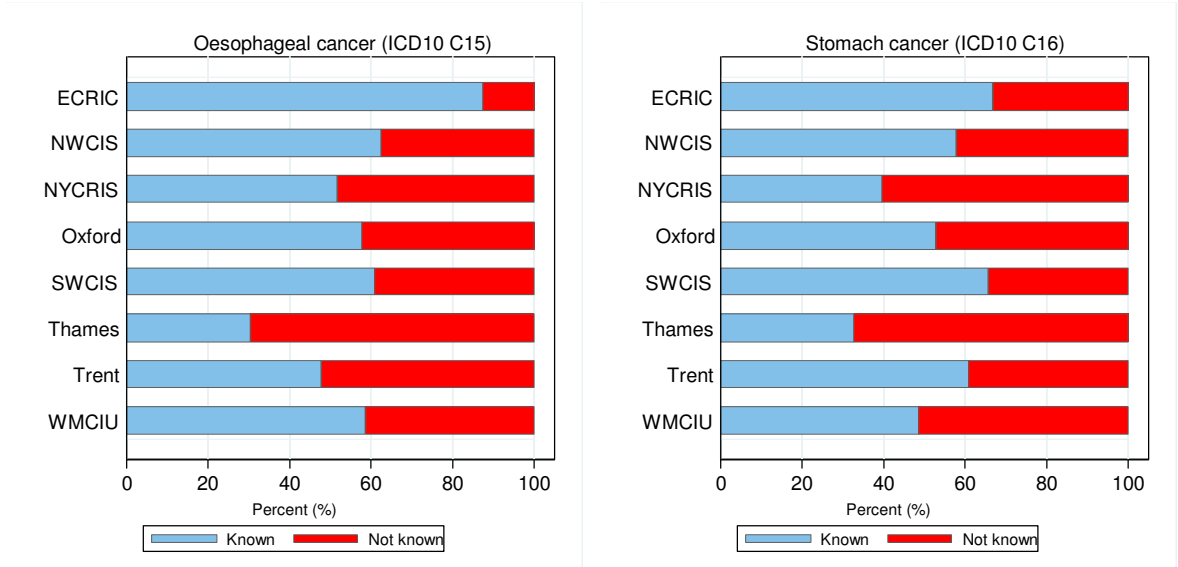


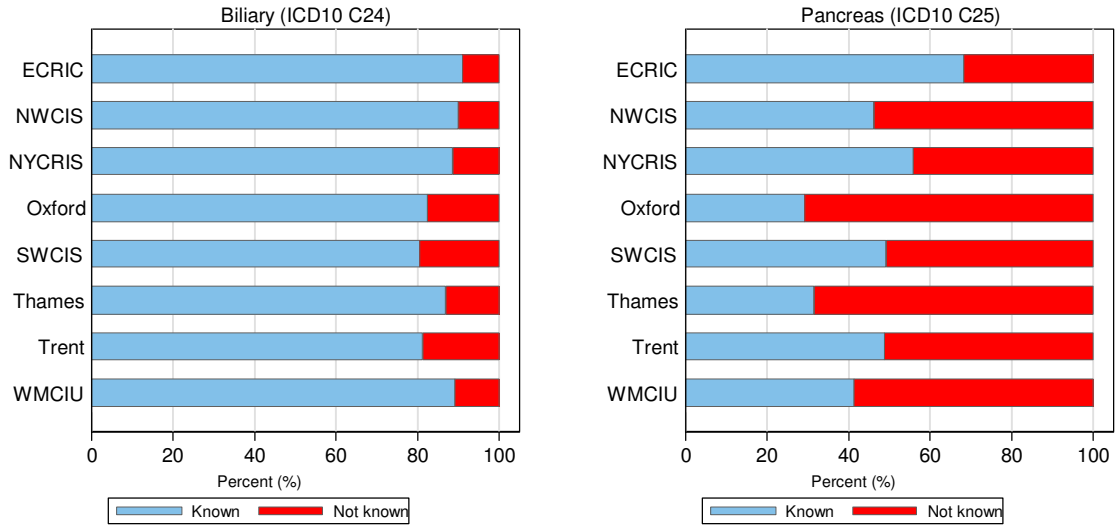
1999-2008



Around 50% of oesophageal, stomach and pancreatic and 85% of biliary cancer registrations had a known anatomical subsite. The proportion of oesophageal cancer registrations with a known anatomical subsite increased between 2006 and 2008. A relatively stable trend was found for the other cancer types.

2008



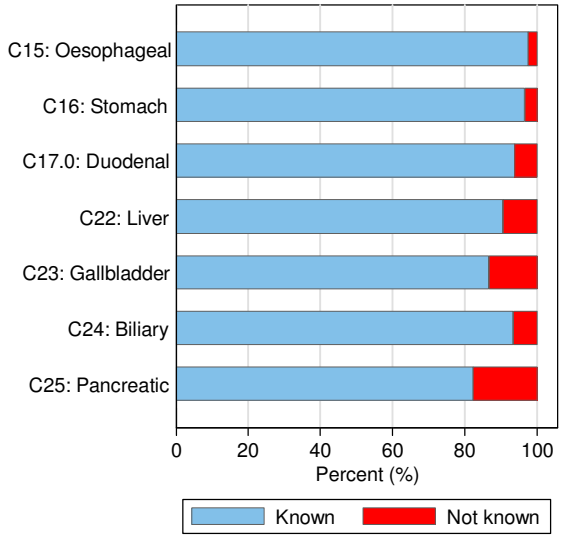


Duodenal, liver and gallbladder cancer were not included in this section. Duodenal cancer is defined by the ICD10 4 digit code of C17.0 (see Appendix 1). Those with an unspecified anatomical location in the C17 (malignant neoplasm of the small intestine) group are defined as C17.8 (overlapping lesion of small intestine) and C17.9 (small intestine, unspecified). In addition to cancers of the duodenum these codes 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 subsite for duodenal cancer were not included in this report. The ICD10 four digit codes for liver cancer are based on morphological definitions and not an anatomical site. Therefore liver cancer was also not included in this section. Finally, all gallbladder cancers are coded as ICD10 C23. There are no further divisions in this group and consequently there are no unspecified anatomical locations.

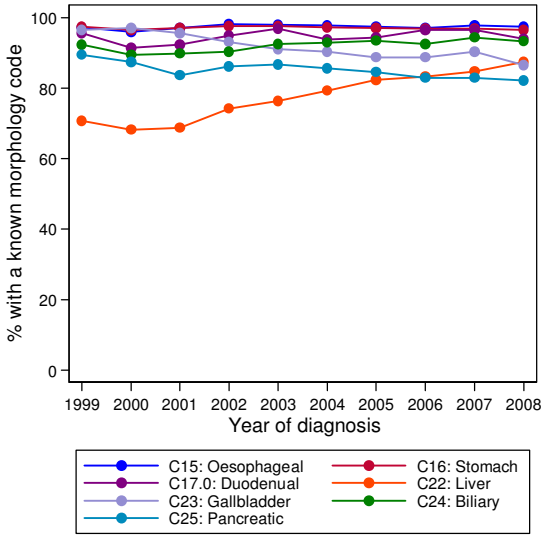
3.5 Morphology

The following graphs show the proportion of registrations with known or not known morphology information for each cancer type. This analysis excludes death certificate only registrations.

2008

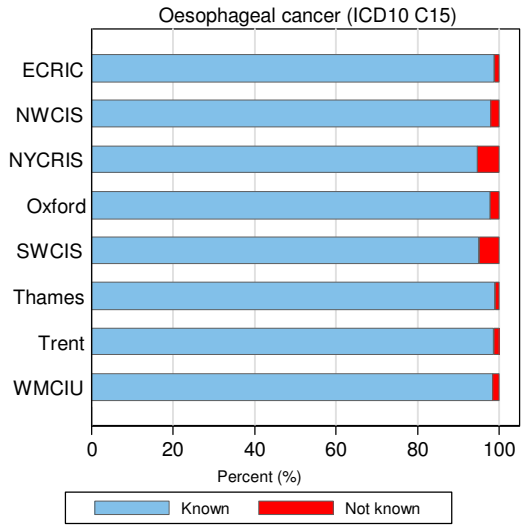


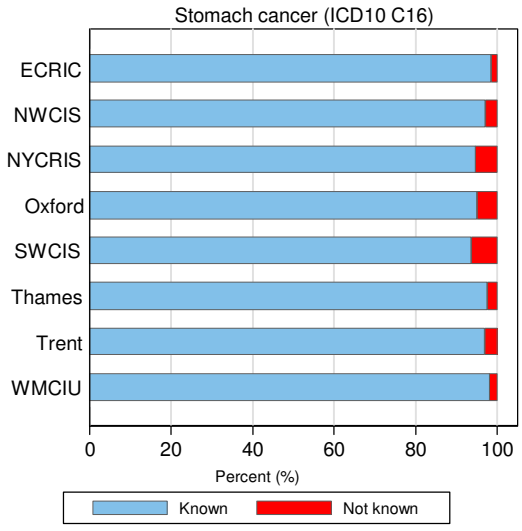
1999-2008

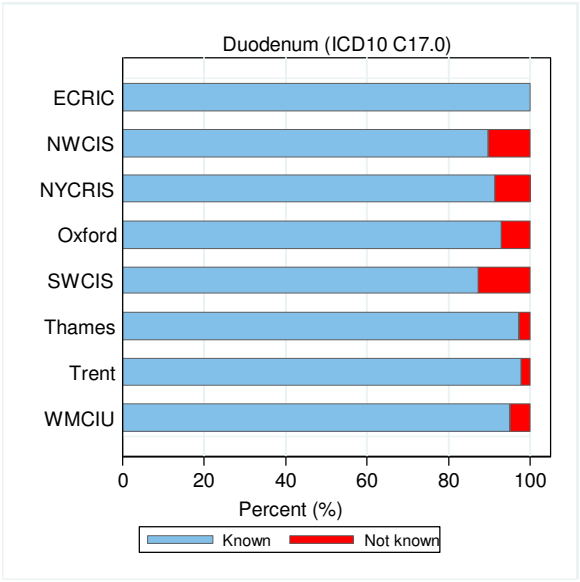


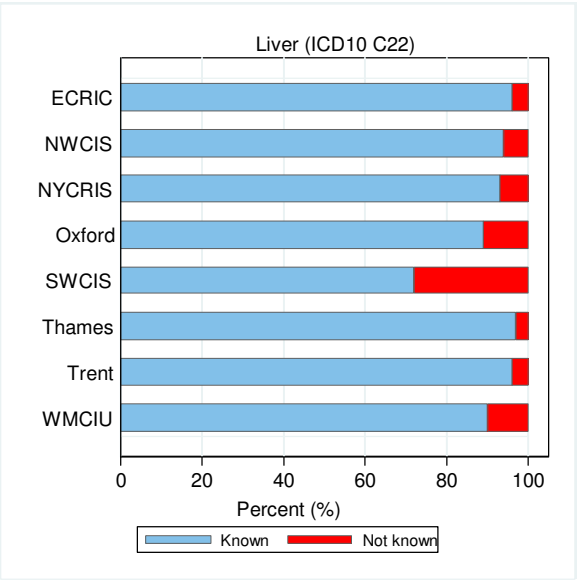
The highest proportion of registrations with known morphology information was found in oesophageal and stomach cancer (over 95%). Pancreatic (82%) and gallbladder (87%) cancer had the lowest proportion with a known morphology. Between 1999 and 2008 there was a relatively stable trend in the proportions of registrations with a known morphology for most of the cancer types, while liver cancer increased from 71% to 87%.

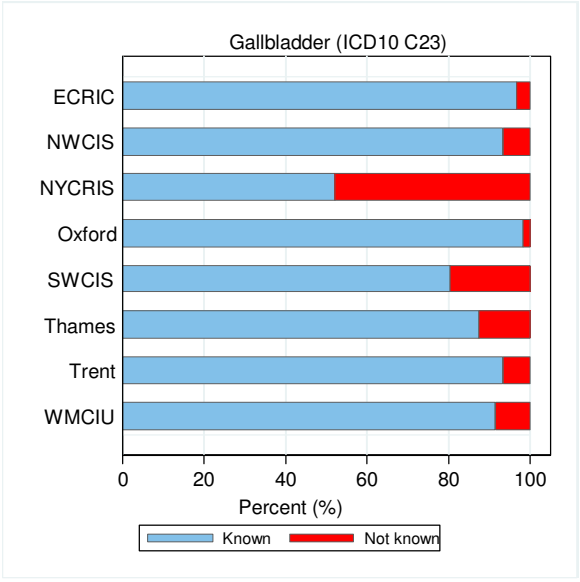
2008

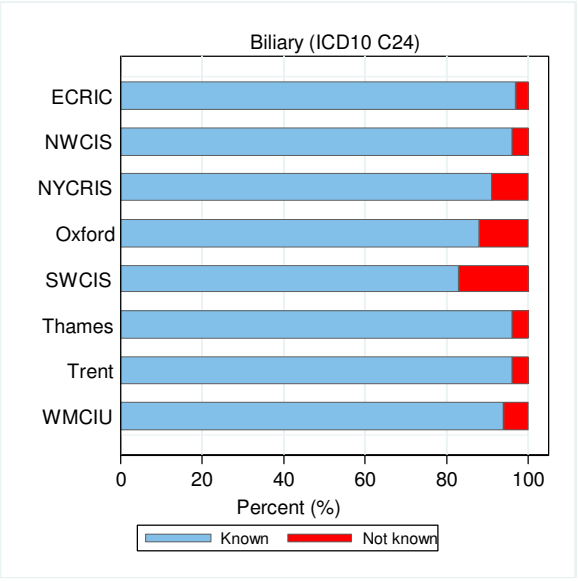


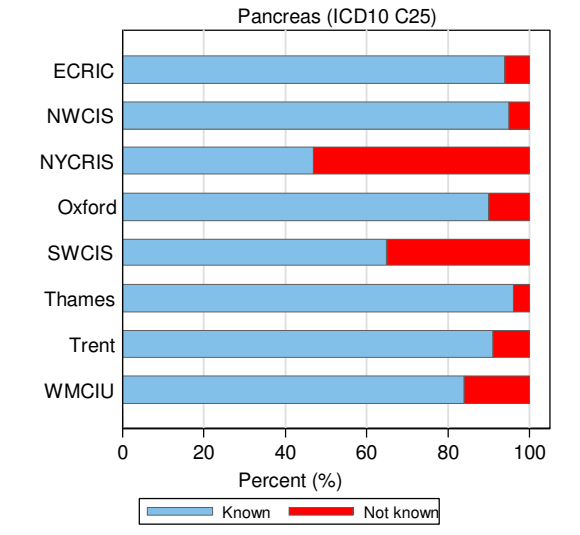










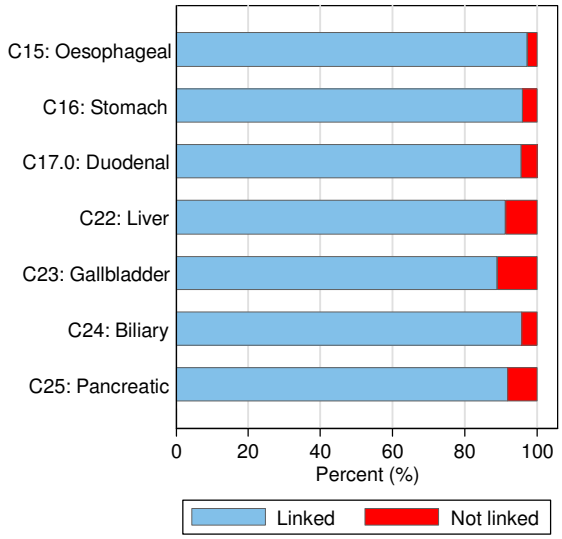


There was variation in the proportion of registrations with a known morphology between cancer registries for gallbladder and pancreatic cancer.

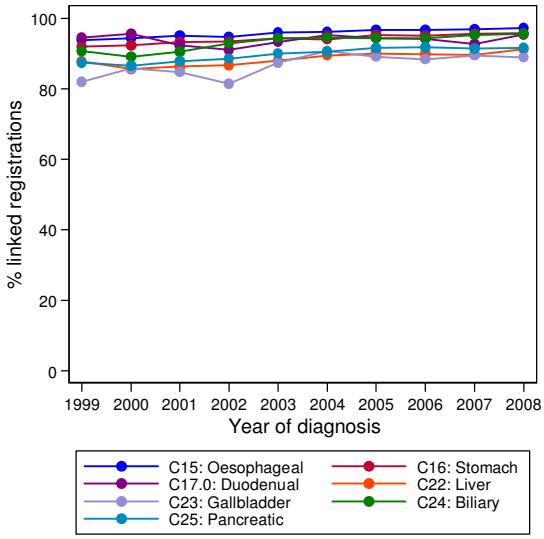
3.6 Linked HES records

The following graphs show the proportion of registrations that were linked and not linked to HES records for each cancer type. This analysis excludes death certificate only registrations.

2008

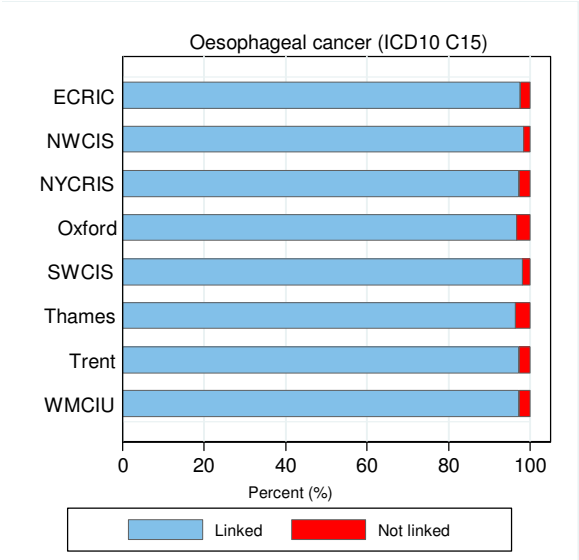


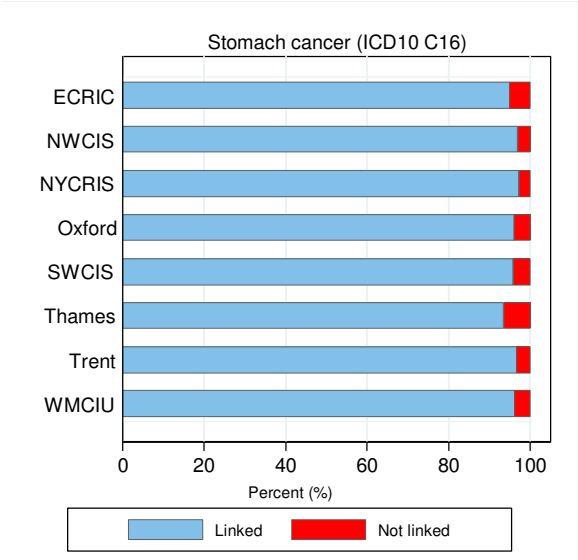
1999-2008

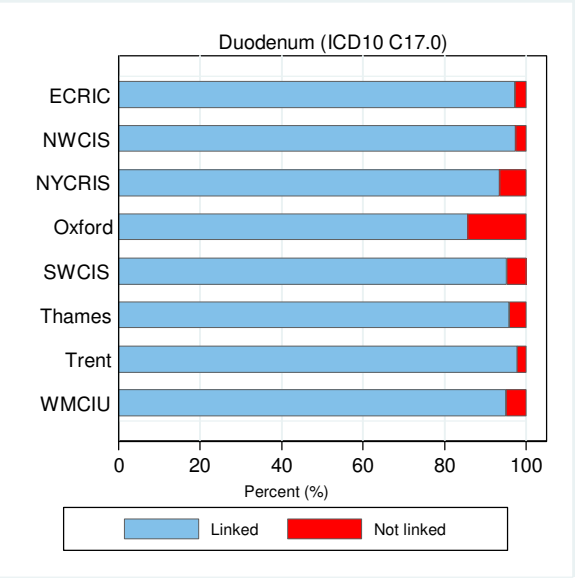


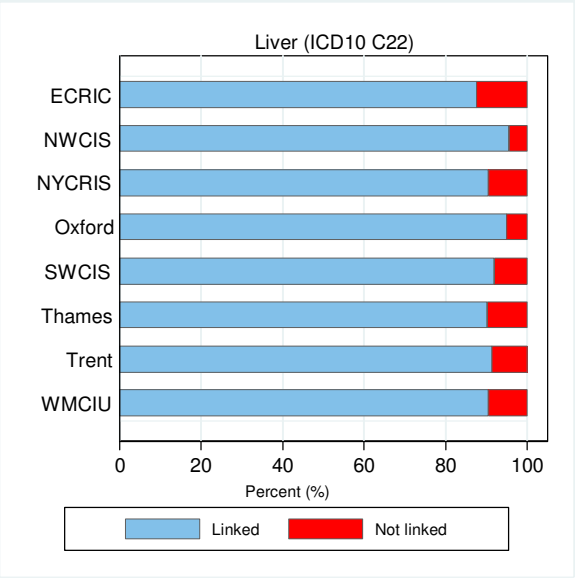
Over 95% of oesophageal, stomach, duodenal and biliary cancer registrations had a linked HES record in 2008. Gallbladder (89%), liver (91%) and pancreatic (92%) cancer had a lower proportion with a matched HES record. Between 1999 and 2008 there was an increase in the proportion of registrations with a linked HES record across all cancer types, even though the proportions were already high in 1999.

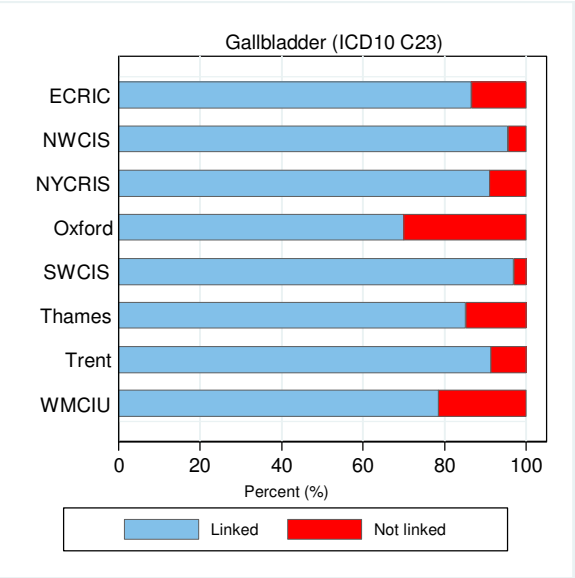
2008

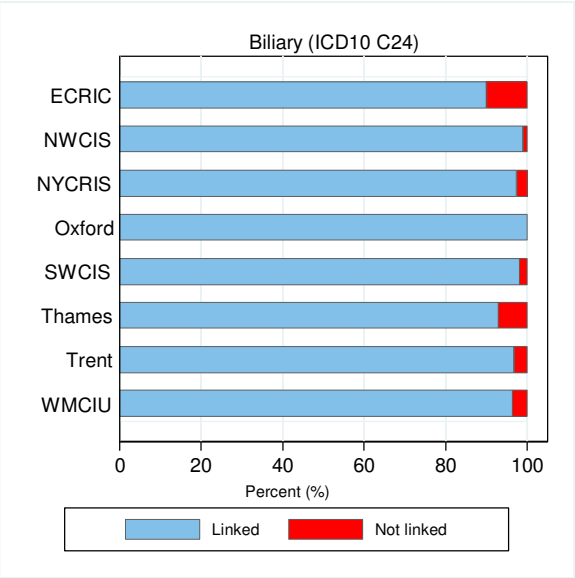


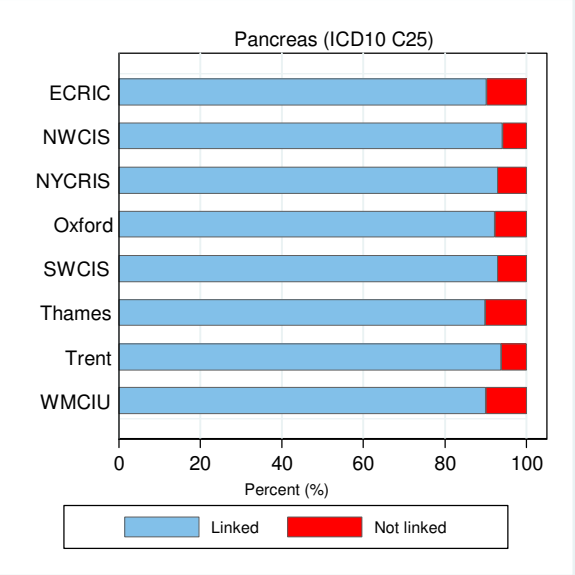










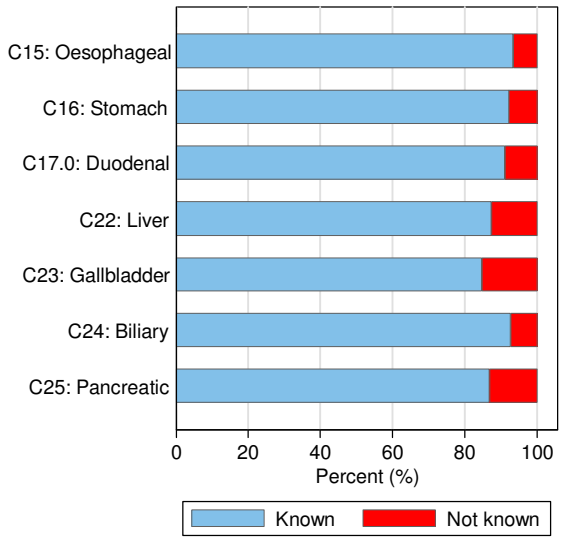


There was variation in the proportion of registrations with a linked HES record between cancer registries for gallbladder cancer.

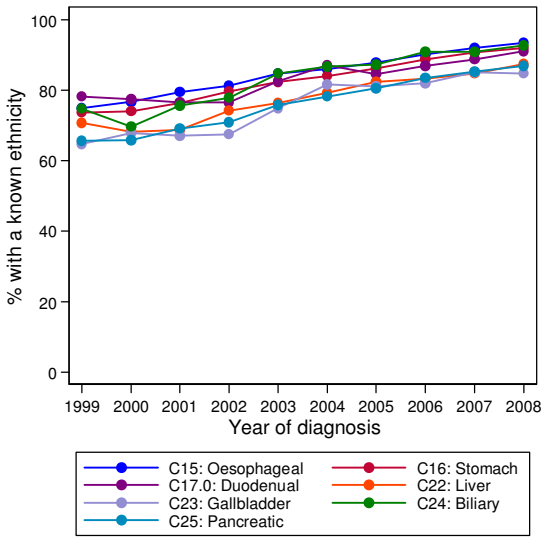
3.7 Ethnicity

The following graphs show the proportion of registrations with a known or not known ethnicity for each cancer type. This analysis excludes death certificate only registrations.

2008

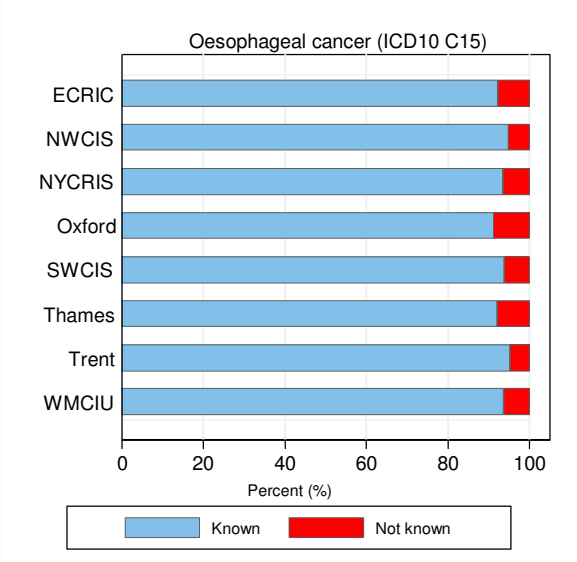


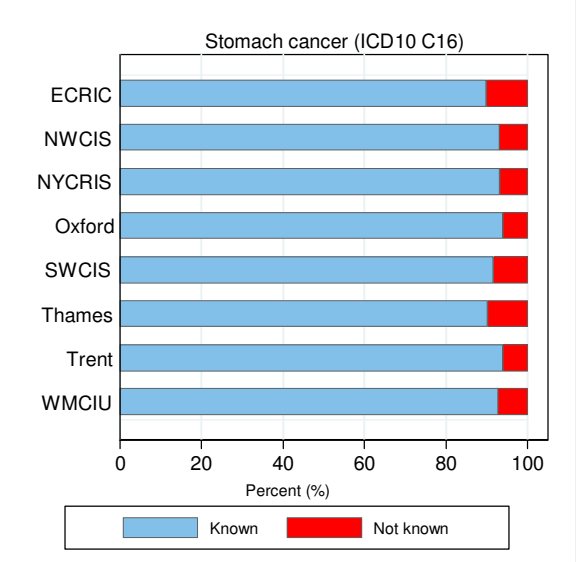
1999-2008

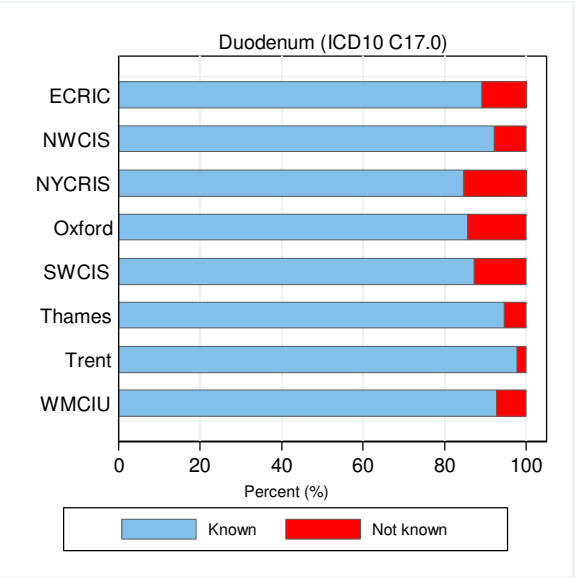


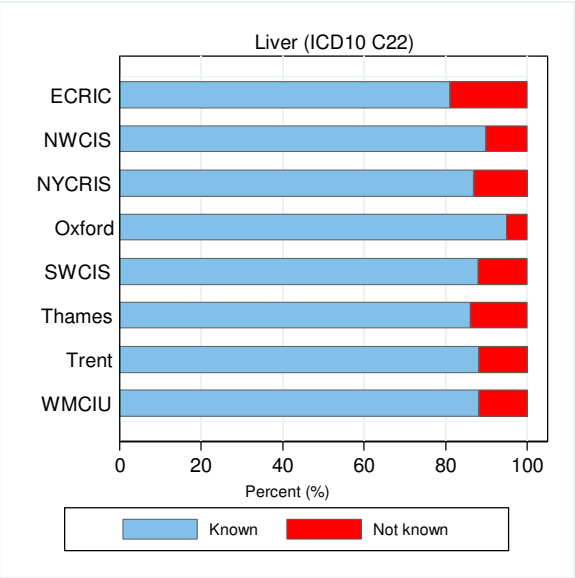
Overall, a high proportion of registrations had a known ethnicity. Gallbladder (85%), pancreas (87%) and liver (87%) cancer had the lowest proportion of registrations with a known ethnicity. Between 1999 and 2008 the proportion of registrations with a known ethnicity increased for all cancer types.

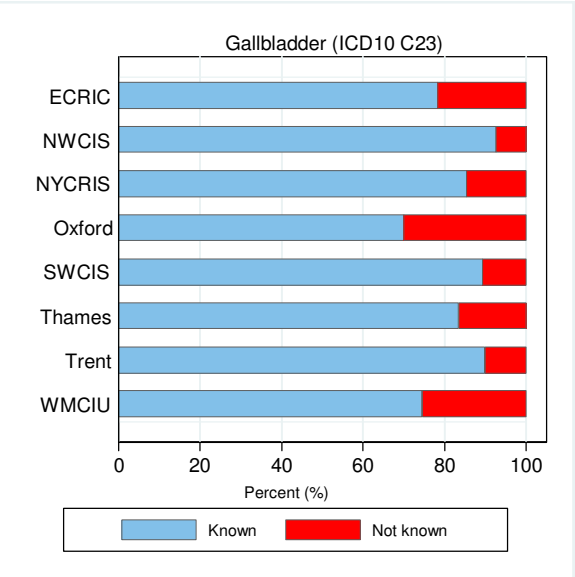
2008

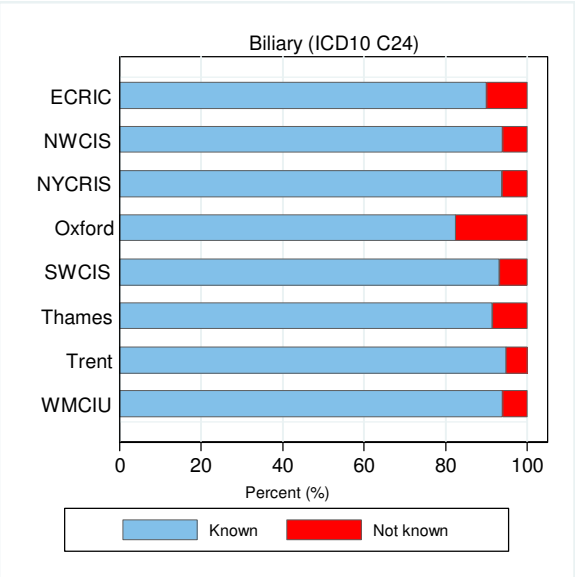


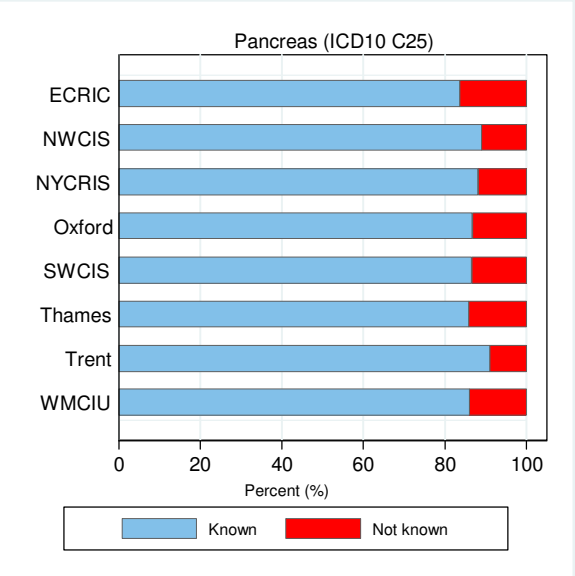












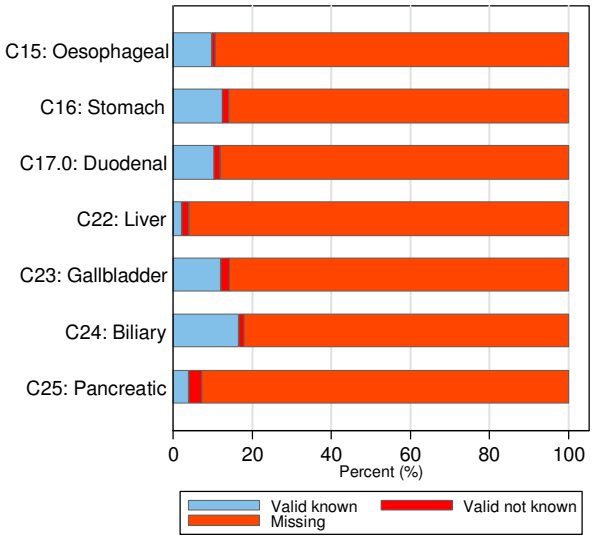
There was variation in the proportion of registrations with a known ethnicity between cancer registries for gallbladder cancer.

This may partly be due to the variation observed in the proportion of registrations with linked HES records.

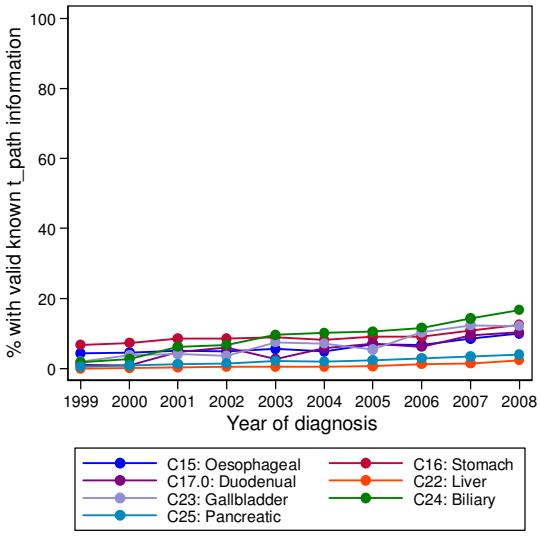
3.8 T stage (pathological)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing T (pathological) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

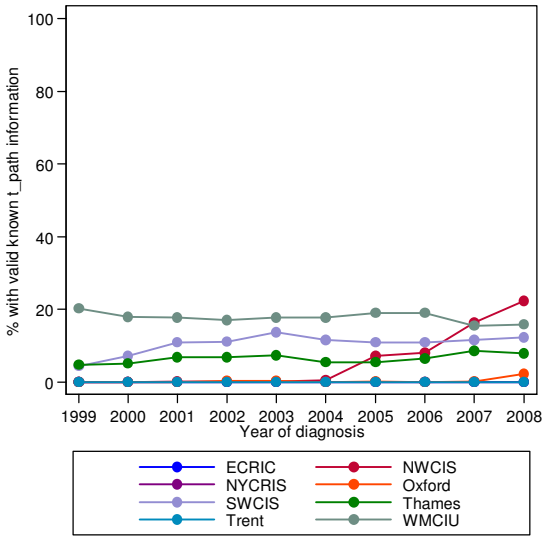


1999-2008



Overall, there were low proportions of T (pathological) stage recorded across each cancer type. Biliary cancer had the highest proportion of registrations with a valid known T stage (17%). Between 1999 and 2008 there was an increase in the proportion of registrations with a valid known T stage for all cancer types.

1999-2008



Not all cancer registries submitted their staging information in the T (pathological) stage field.

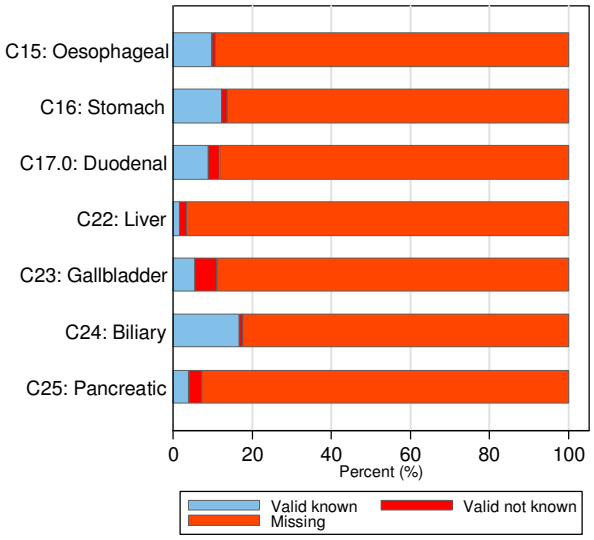
Between 1999 and 2008 there was a relatively stable trend in the proportion of registrations with a valid known T (pathological) stage across most cancer registries.

Since 2004, the proportion of registrations with a valid known T (pathological) stage increased in NWCIS.

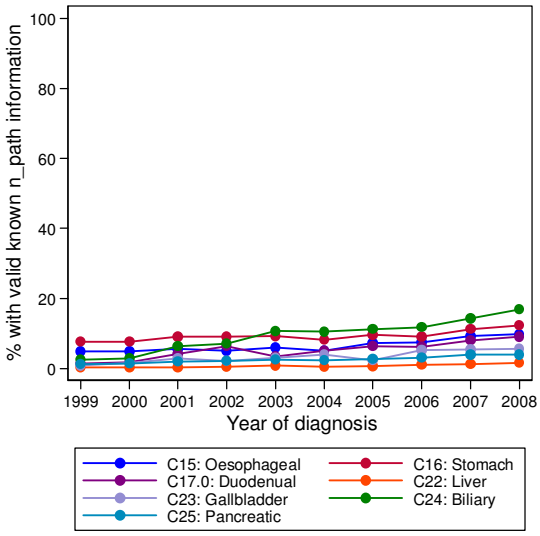
3.9 N stage (pathological)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing N (pathological) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

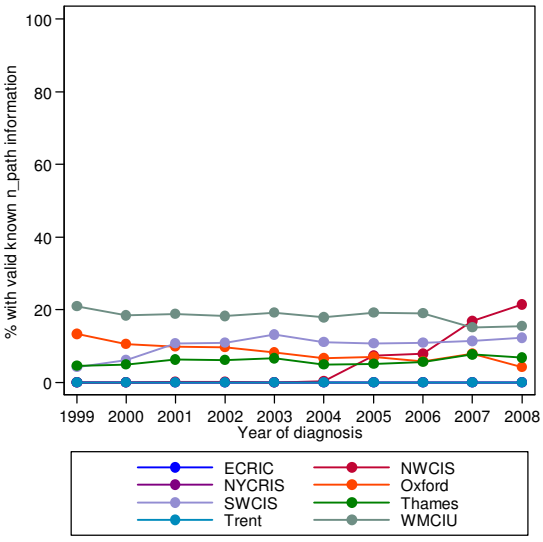


1999-2008



Overall, there were low proportions of N (pathological) stage recorded across each cancer type. Biliary cancer had the highest proportion of registrations with a valid known N stage (17%). Between 1999 and 2008 there was an increase in the proportion of registrations with a valid known N stage for most cancer types.

1999-2008



Not all cancer registries submitted their staging information in the N (pathological) stage field.

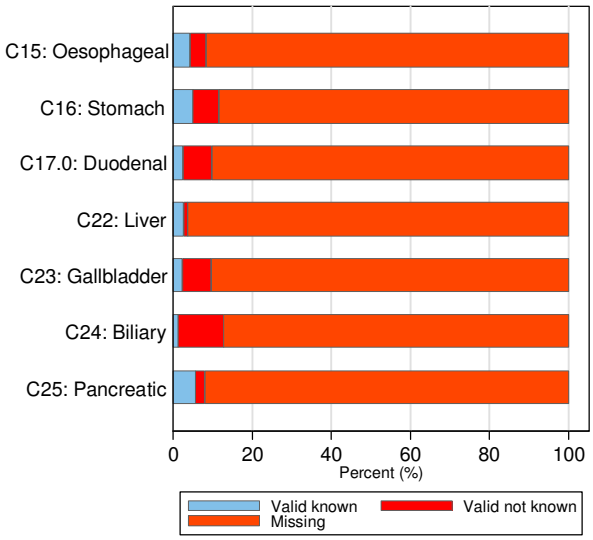
Between 1999 and 2008 there was a relatively stable trend in the proportion of registrations with a valid known N (pathological) stage across most cancer registries.

Since 2004, the proportion of registrations with a valid known N (pathological) stage increased in NWCIS.

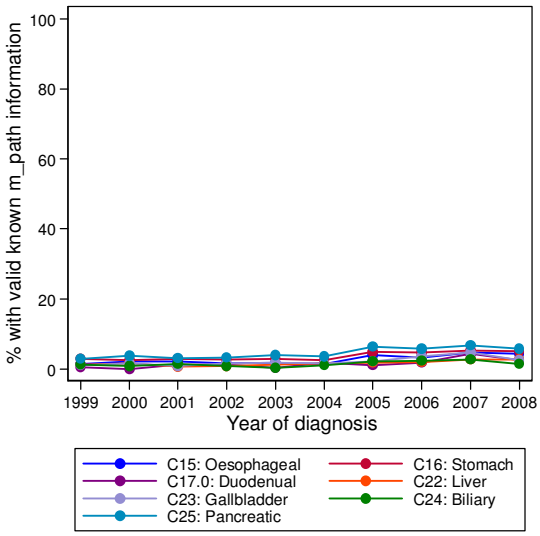
3.10 M stage (pathological)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing M (pathological) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

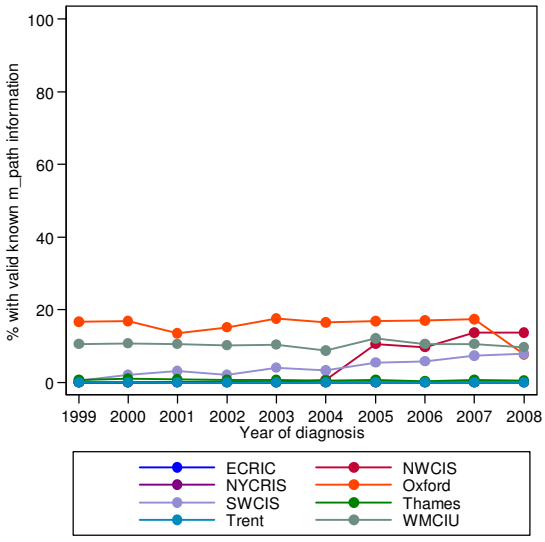


1999-2008



Overall, there were low proportions of M (pathological) stage recorded across each cancer type. Pancreatic cancer had the highest proportion of registrations with a valid known M stage (6%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known M stage for some cancer types. There were higher proportions of valid not known M (pathological) stage in some cancer types.

1999-2008



Not all cancer registries submitted their staging information in the M (pathological) stage field.

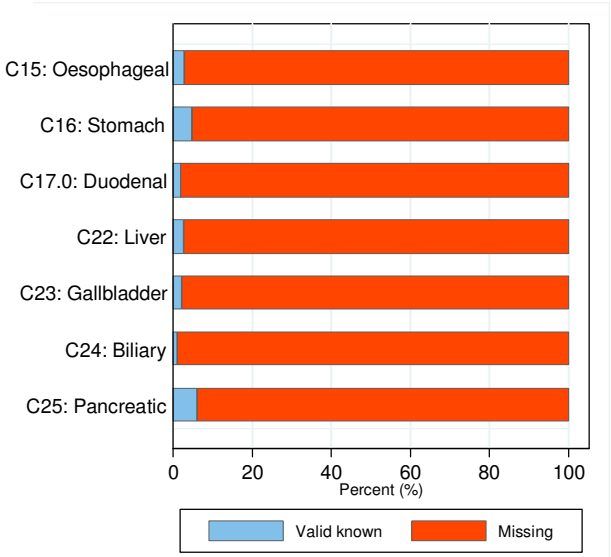
Between 1999 and 2008 there was a relatively stable trend in the proportion of registrations with a valid known M (pathological) stage across most cancer registries.

Since 2004, the proportion of registrations with a valid known M (pathological) stage increased in NWCIS.

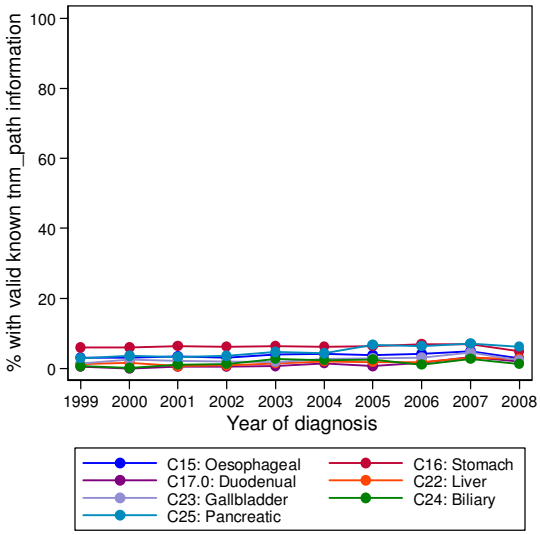
3.11 TNM stage (pathological)

The following graphs show the proportion of registrations with a valid known or a missing TNM (pathological) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

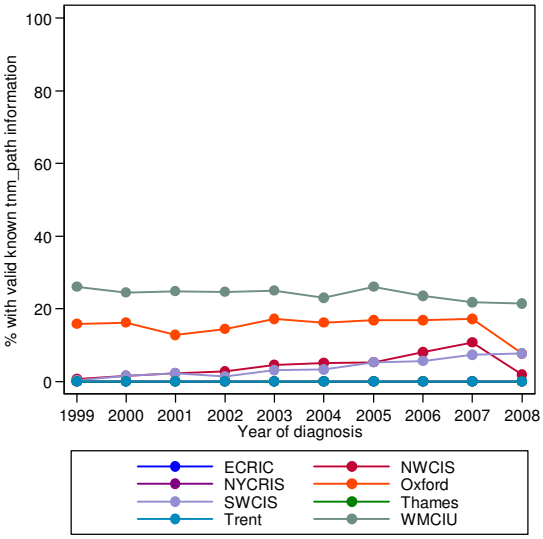


1999-2008



Overall, there were low proportions of TNM (pathological) stage recorded across each cancer type. Pancreatic cancer had the highest proportion of registrations with a valid known TNM stage (6%). Between 1999 and 2008 there was a relatively stable trend in the proportion of registrations with a valid known TNM stage for most cancer types.

1999-2008



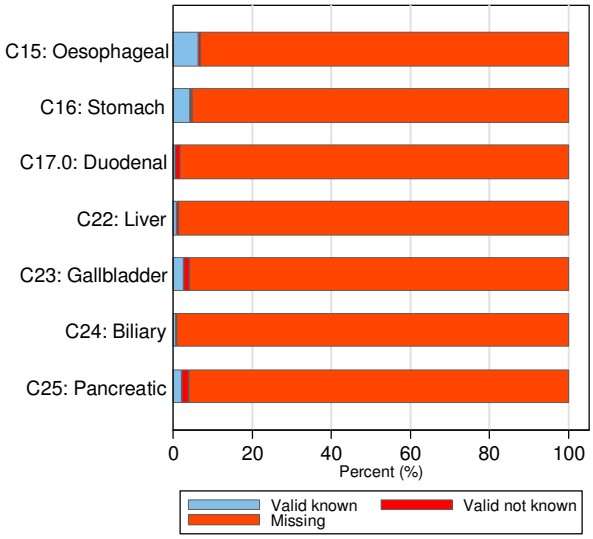
Not all cancer registries submitted their staging information in the TMN (pathological) stage field.

Between 1999 and 2008 there was a relatively stable trend in the proportion of registrations with a valid known TNM (pathological) stage across most cancer registries.

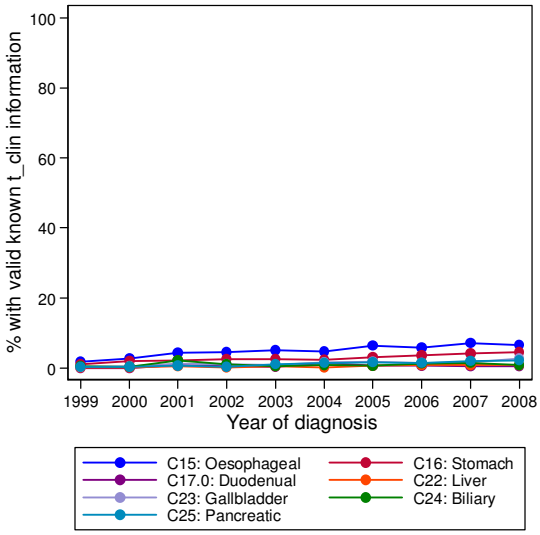
3.12 T stage (clinical)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing T (clinical) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

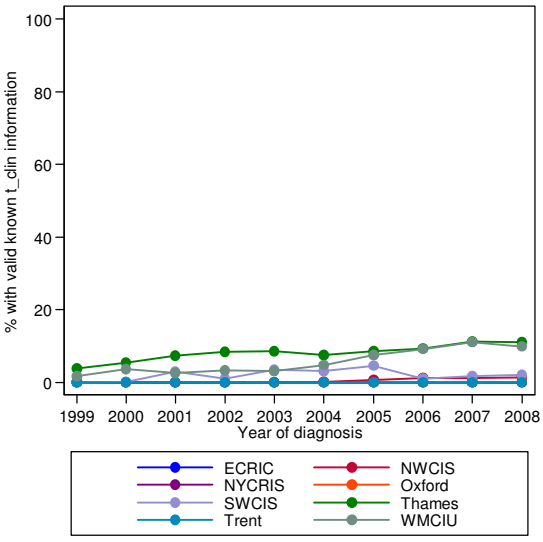


1999-2008



Overall, there were low proportions of T (clinical) stage recorded across each cancer type. Oesophageal cancer had the highest proportion of registrations with a valid known T stage (6%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known T stage for some cancer types.

1999-2008



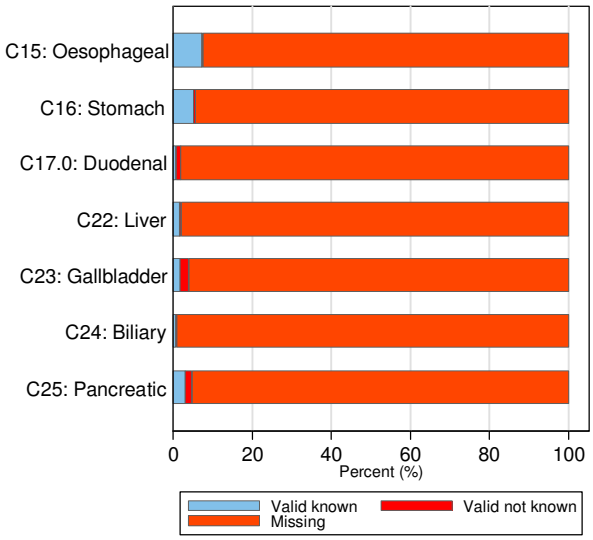
Not all cancer registries submitted their staging information in the T (clinical) stage field.

Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known T (clinical) stage in Thames and WMCIU.

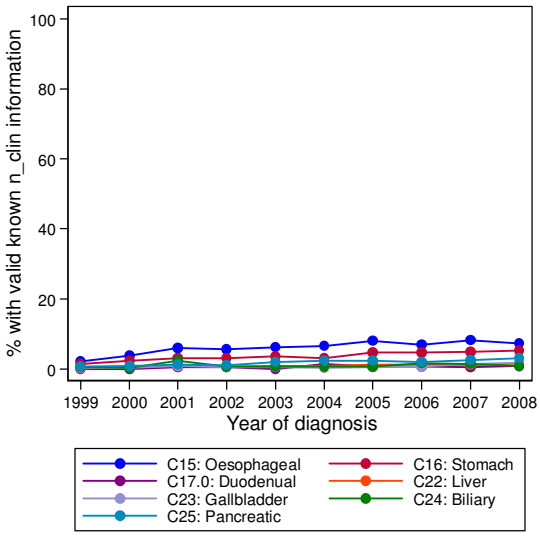
3.13 N stage (clinical)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing N (clinical) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

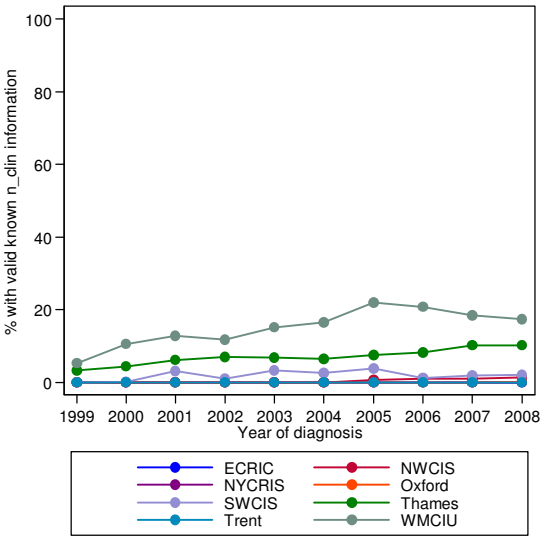


1999-2008



Overall, there were low proportions of N (clinical) stage recorded across each cancer type. Oesophageal cancer had the highest proportion of registrations with a valid known N stage (7%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known N stage for most cancer types.

1999-2008



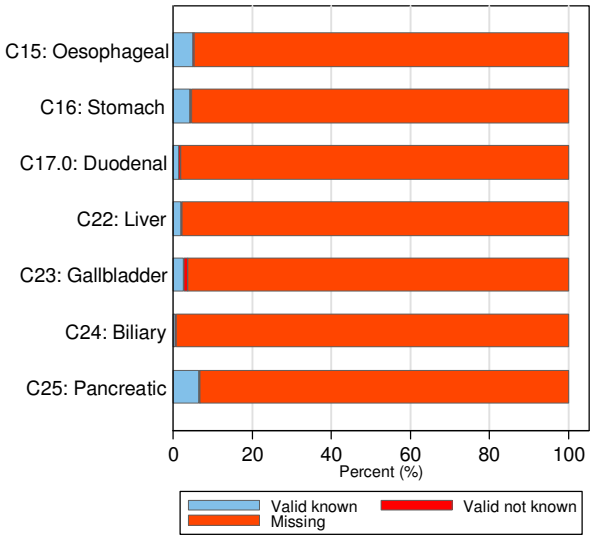
Not all cancer registries submitted their staging information in the N (clinical) stage field.

Between 1999 and 2008 there was an increase in the proportion of registrations with a valid known N (clinical) stage in Thames and WMCIU.

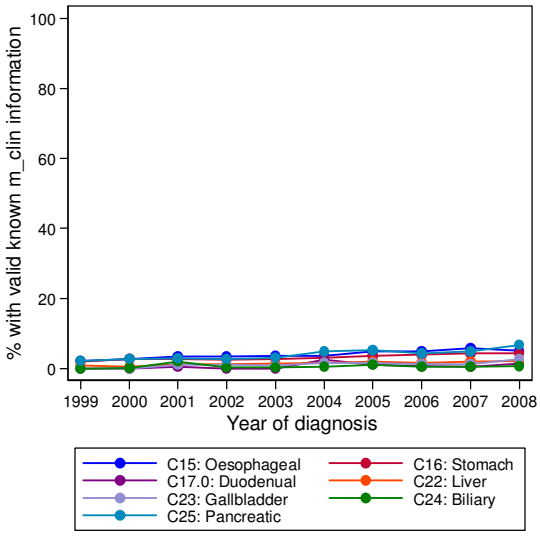
3.14 M stage (clinical)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing M (clinical) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

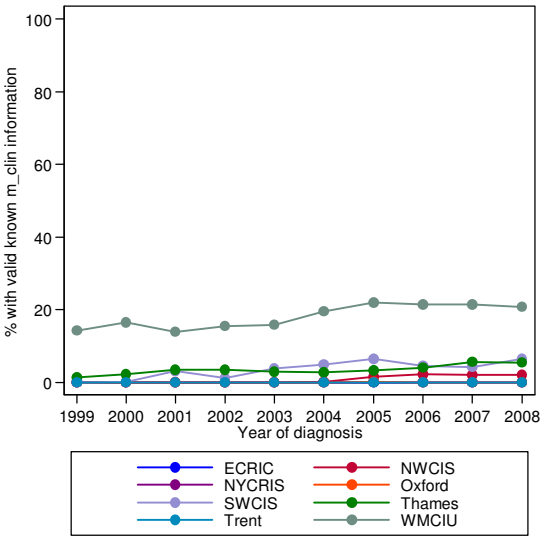


1999-2008



Overall, there were low proportions of M (clinical) stage recorded across each cancer type. Pancreatic cancer had the highest proportion of registrations with a valid known M stage (7%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known M stage for most cancer types.

1999-2008



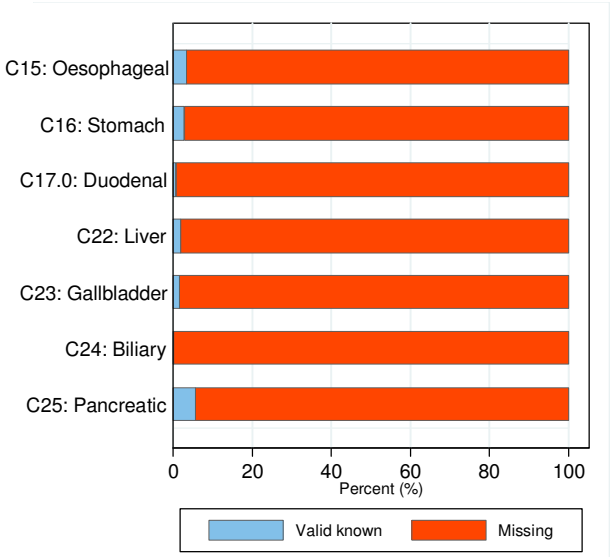
Not all cancer registries submitted their staging information in the M (clinical) stage field.

Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known M (clinical) stage in WMCIU.

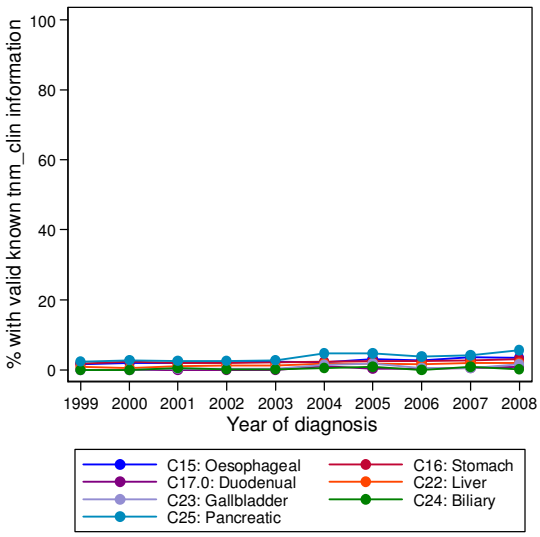
3.15 TNM stage (clinical)

The following graphs show the proportion of registrations with a valid known or a missing TNM (clinical) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

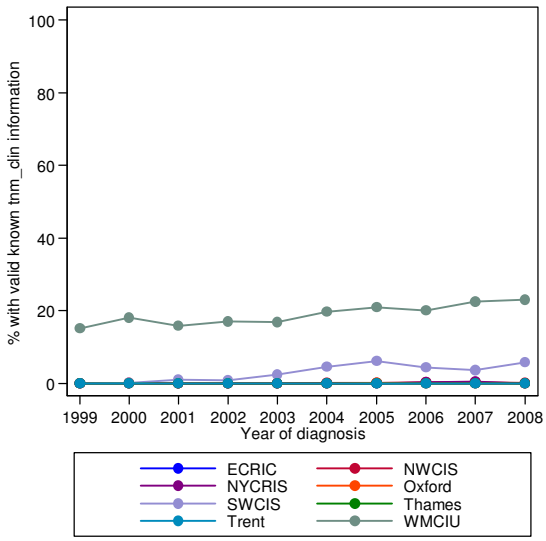


1999-2008



Overall, there were very low proportions of TNM (clinical) stage recorded across each cancer type. Pancreatic cancer had the highest proportion of registrations with a valid known TNM stage (6%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known TNM stage for some cancer types.

1999-2008



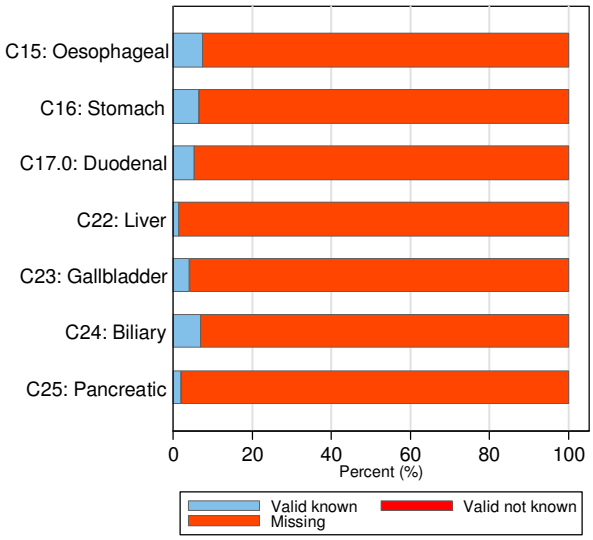
Not all cancer registries submitted their staging information in the TNM (clinical) stage field.

Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known TNM (clinical) stage in WMCIU.

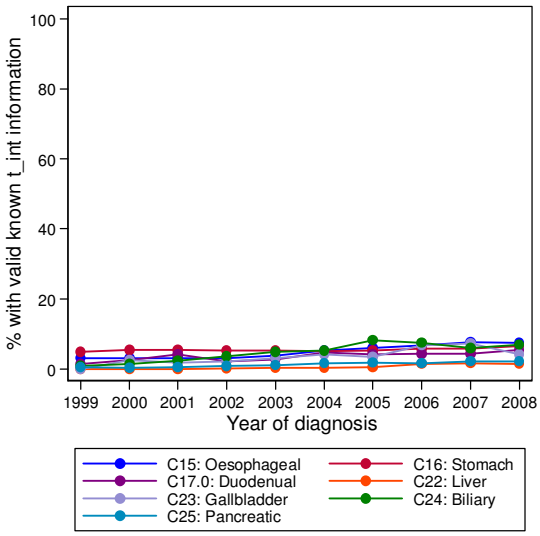
3.16 T stage (integrated)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing T (integrated) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

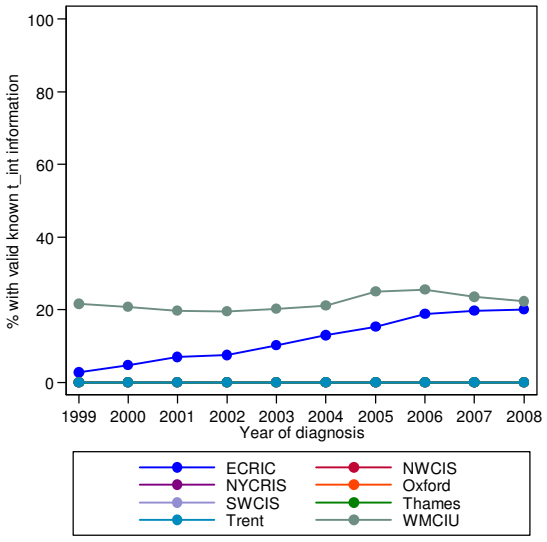


1999-2008



Overall, there were low proportions of T (integrated) stage recorded across each cancer type. Oesophageal, stomach and biliary cancer had the highest proportions of registrations with a valid known T stage (7%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known T stage for some cancer types.

1999-2008



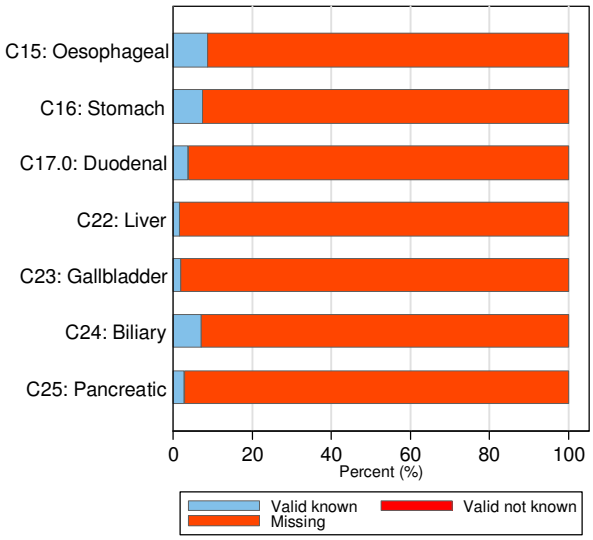
Only two cancer registries (ECRIC and WMCIU) submitted their staging information using the T (integrated) stage field.

Between 1999 and 2008 there was an increase in the proportion of registrations with a valid known T (integrated) stage in ECRIC.

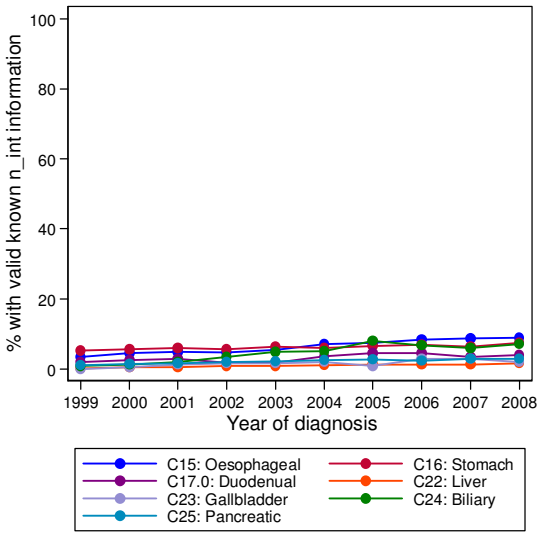
3.17 N stage (integrated)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing N (integrated) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

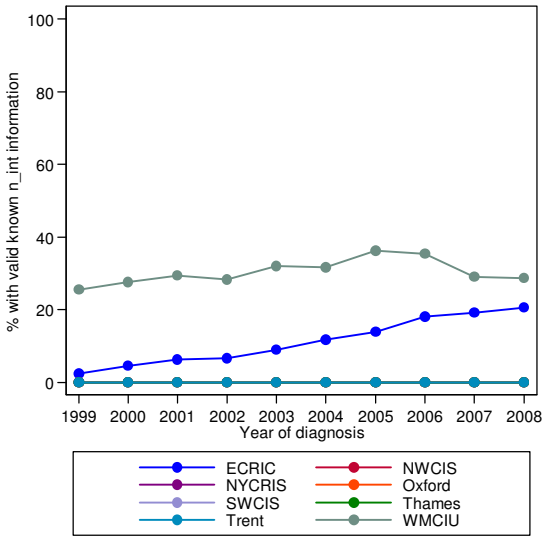


1999-2008



Overall, there were low proportions of N (integrated) stage recorded across each cancer type. Oesophageal cancer had the highest proportion of registrations with a valid known N stage (9%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known N stage for some cancer types.

1999-2008



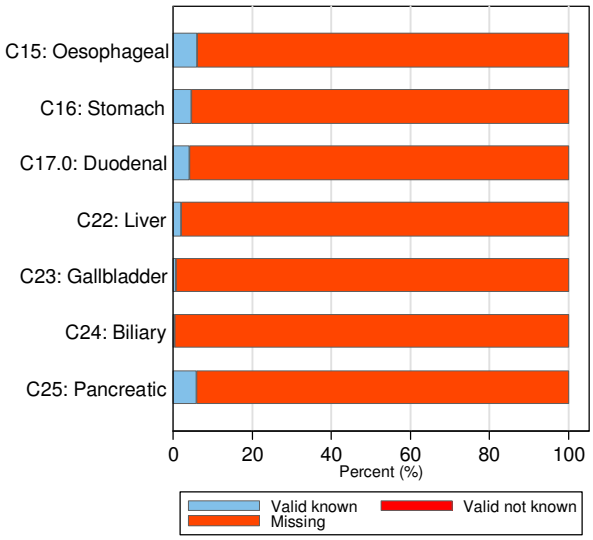
Only two cancer registries (ECRIC and WMCIU) submitted their staging information using the N (integrated) stage field.

Between 1999 and 2008 there was an increase in the proportion of registrations with a valid known N (integrated) stage in ECRIC.

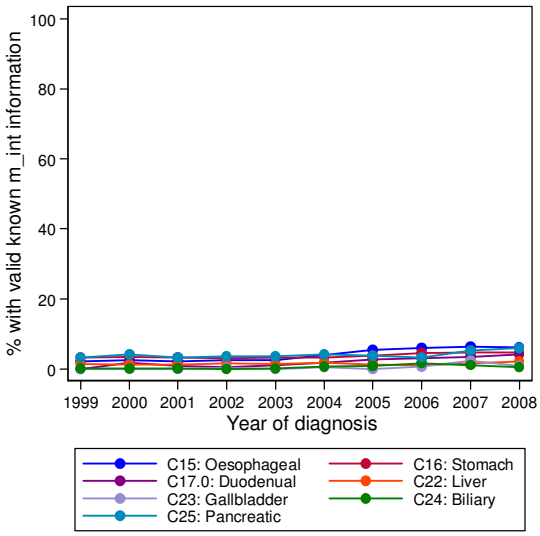
3.18 M stage (integrated)

The following graphs show the proportion of registrations with a valid known, a valid not known or missing M (integrated) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

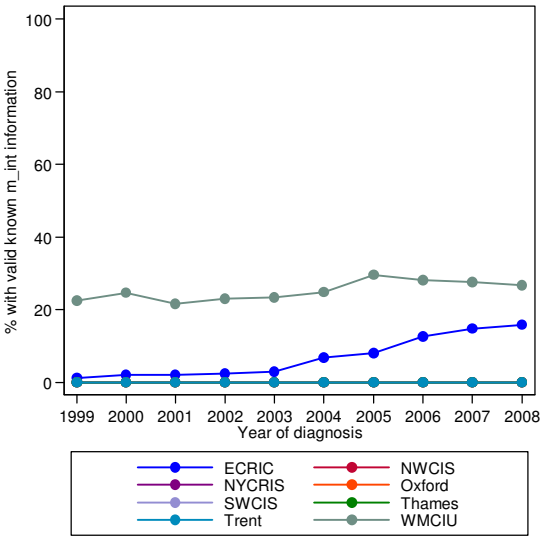


1999-2008



Overall, there were low proportions of M (integrated) stage recorded across each cancer type. Oesophageal and pancreatic cancers had the highest proportion of registrations with a valid known M stage (6%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known M stage for some cancer types.

1999-2008



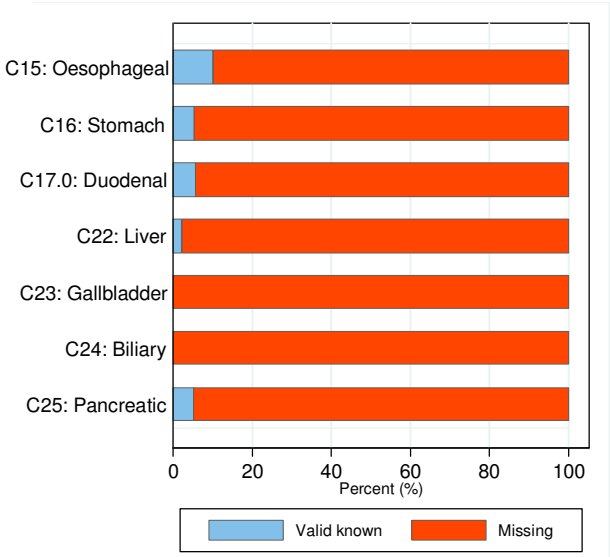
Only two cancer registries (ECRIC and WMCIU) submitted their staging information using the M (integrated) stage field.

Between 2003 and 2008 there was an increase in the proportion of registrations with a valid known M (integrated) stage in ECRIC.

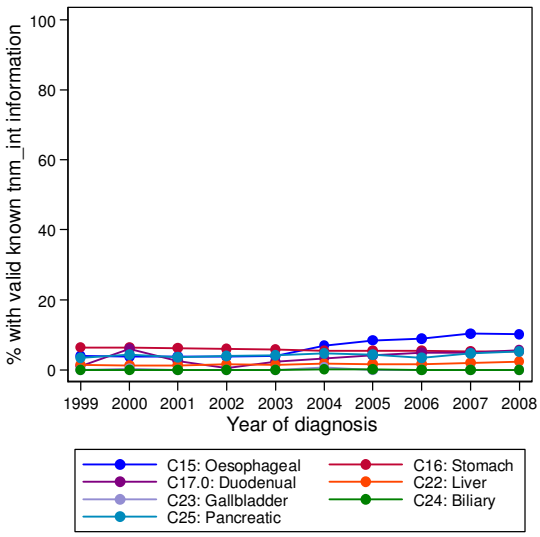
3.19 TNM stage (integrated)

The following graphs show the proportion of registrations with a valid known or a missing TNM (integrated) stage for each cancer type. This analysis excludes death certificate only registrations.

2008

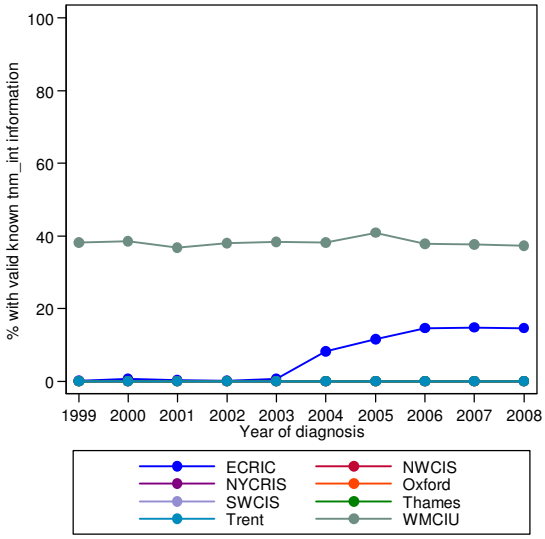


1999-2008



Overall, there were low proportions of TNM (integrated) stage recorded across each cancer type. Oesophageal cancer had the highest proportion of registrations with a valid known TNM stage (10%). Between 1999 and 2008 there was a slight increase in the proportion of registrations with a valid known TNM stage for some cancer types.

1999-2008



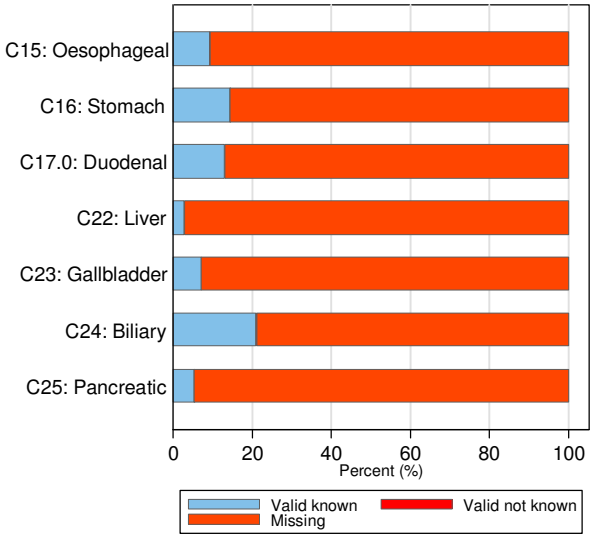
Only two cancer registries (ECRIC and WMCIU) submitted their staging information using the TNM (integrated) stage field.

Between 2003 and 2008 there was an increase in the proportion of registrations with a valid known TNM (integrated) stage in ECRIC.

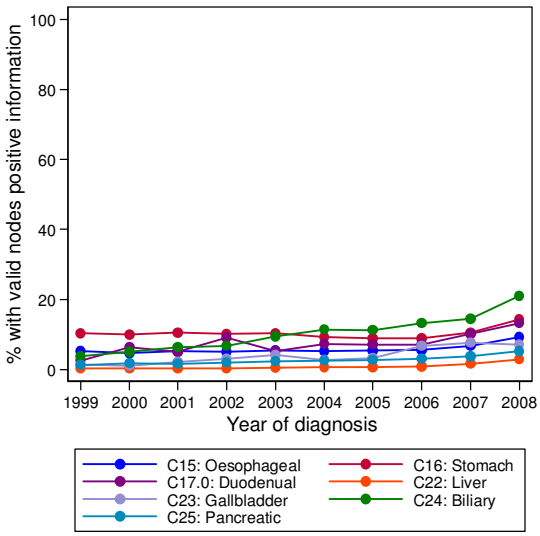
3.20 Nodes positive

The following graphs show the proportion of registrations with valid known, valid not known and missing nodes positive information for each cancer type. This analysis excludes death certificate only registrations.

2008

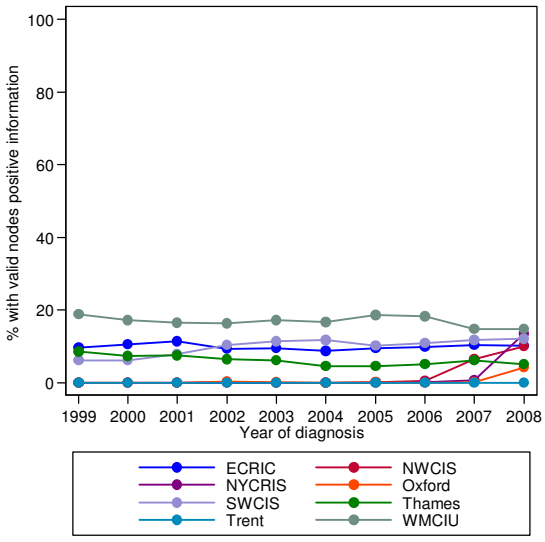


1998-2008



Overall, there were low proportions of nodes positive information recorded across each cancer type. Biliary cancer had the highest proportion of registrations with valid known nodes positive information (21%). Between 1999 and 2008 there was an increase in the proportion of registrations with valid known nodes positive information for all cancer types.

1999-2008



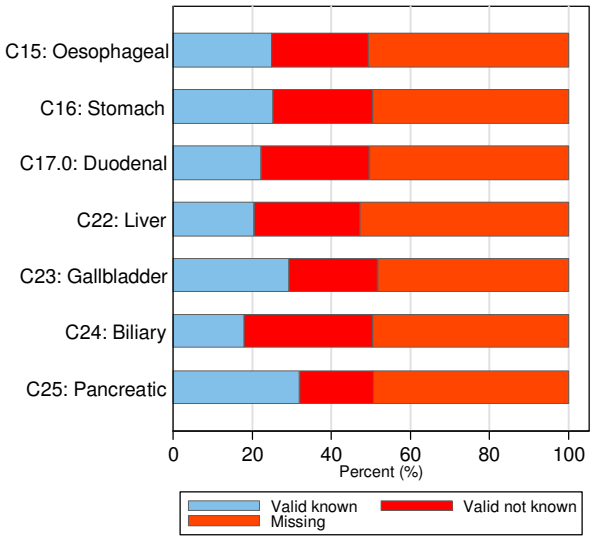
Between 1999 and 2008 there was a relatively stable trend in the proportion of registrations with valid known nodes positive information across most cancer registries.

Some cancer registries have only recorded this information in the most recent years.

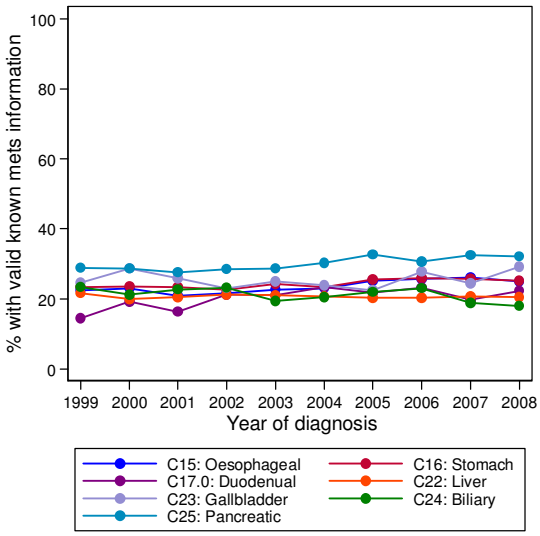
3.21 Distant metastases

The following graphs show the proportion of registrations with valid known, valid not known and missing distant metastases information for each cancer type. This analysis excludes death certificate only registrations.

2008

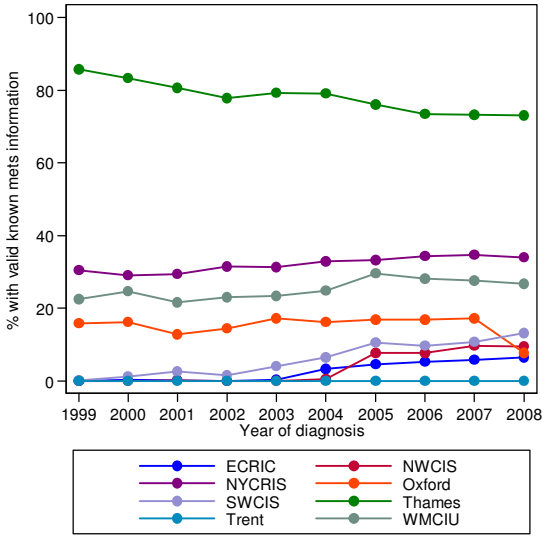


1998-2008



Overall, there were low proportions of distant metastases recorded across each cancer type. Pancreatic cancer had the highest proportion of registrations with valid known metastases information (32%). Around 50% of registrations were either valid known and valid not known across all cancer types. Between 1999 and 2008 there was a stable trend in the proportion of registrations with valid known metastases information.

1999-2008



Thames had the highest proportion of registrations with valid known metastases information.

There have been increases in the proportion of registrations with valid known metastases information in ECRIC, NWCIS and SWCIS.

3.22 Completeness

Overall, the estimated completeness of the OG and HPB cancer datasets was good. In 2008, less than 1% of patients with all cancer types were estimated to have been missed by the cancer registration process (Table 2).

Table 2: Completeness of the OG and HPB cancer dataset

Cancer type (ICD10 code)	Cancer registry dataset	HES-onlys	(%)
Oesophageal cancer (C15)	6,671	17	0.3
Stomach cancer (C16)	6,124	23	0.4
Duodenal cancer (C17.0)	367	1	0.3
Liver cancer (C22)	2,978	20	0.7
Gallbladder cancer (C23)	551	4	0.7
Biliary cancer (C24)	777	0	0.0
Pancreas cancer (C25)	6,817	44	0.6

The completeness of the datasets was calculated by extracting HES only records with a cancer diagnosis and relevant surgical procedure. The combination of these codes increased the certainty that these patients were true cancer cases and not just a record of a suspicion of cancer. However, a low proportion of patients with some of these cancer types will have surgery. Therefore as this method relies on a record of surgery to identify cancer cases it may over-estimate the completeness of the ascertainment of these cancers.

4. Key findings

- The proportion of death certificate only registrations ranged from 1.3% (biliary cancer) to 7.0% (liver cancer). Primary liver and pancreatic cancer had the highest proportions of DCO registrations. The proportion of DCO registrations was decreasing over time (1999-2008).
- The proprtions of microscopically verified cases ranged from 46.5% (liver cancer) to 93.1% (oesophageal cancer). Smaller proportions of registrations were microscopically verified in primary liver (46.5%) and pancreatic (47.9%) cancers compared to oesophageal and stomach cancers. Over half of liver (52.0%) and pancreatic (51.1%) cancers were only clinically verified.
- Only around half of oesophageal (55.6%), stomach (51.4%) and pancreatic (46.8%) cancer registrations had a known anatomical subsite, while this figure was 86.7% for biliary cancer.
- Over 90% of oesophageal, stomach, duodenal, liver and biliary cancer registrations had known morphology information.
- Over 89% of all OG and HPB cancers had a linked HES record. The proportion of HES linked records increased over time.
- Over 90% of oesophageal (93.5%), stomach (92.1%), duodenal (91.1%) and biliary (92.7%) cancer registrations had a known ethnicity. The proportion of registrations with a known ethnicity increased over the ten-year period.
- The availability of information from all the staging fields studied (TNM, mets and nodes postive) was poor, although in some cases there was an increase in the proportion with a valid known record over time.
- In 2008, only small proportions of OG and HPB cancer patients were estimated to have been missed by the cancer registration process, using a method which identifies HES only records with both a diagnosis and relevant procedure code.

5. Conclusions

This report has investigated the data quality of the registrations held within the NCDR upper gastrointestinal cancer dataset.

The proportion of death certificate only registrations was generally low and was decreasing over the ten-year period 1999-2008. These registrations would have to be excluded from survival analysis and may indicate incomplete case ascertainment, both factors which could potentially bias the survival estimates. It is important that work continues to reduce the proportion of these registrations.

The proportion of registrations with a valid ethnic group classification was high and has increased over time. Also, a high proportion of all cancer types had a linked record in HES. Again, this increased over the study period. These increasing trends are likely to continue alongside improvements in the linkage between the two datasets.

Overall, the availability of staging information was poor and this should be improved. However, it is encouraging to note that in general the proportion of registrations with valid known staging information is increasing over time. Various national projects have been developed to improve the availability of staging information, so with time this may improve.

This report also shows that better anatomical and morphological 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 missed registrations.

Appendix 1: 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 2: List of ICD10 codes and procedure codes used in the completeness analysis.

Oesophageal cancer (ICD10 C15)	G011	Oesophagogastrectomy and anastomosis of oesophagus to stomach
	G012	Oesophagogastrectomy and anastomosis of oesophagus to transposed jejunum
	G013	Oesophagogastrectomy and anastomosis of oesophagus to jejunum NEC
	G018	Other specified excision of oesophagus and stomach
	G019	Unspecified excision of oesophagus and stomach
	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
	G028	Other specified total excision of oesophagus
	G029	Unspecified total excision of oesophagus
	G031	Partial oesophagectomy and end to end anastomosis of oesophagus
	G032	Partial oesophagectomy and interposition of microvascularly attached jejunum
	G033	Partial oesophagectomy and anastomosis of oesophagus to transposed jejunum
	G034	Partial oesophagectomy and anastomosis of oesophagus to jejunum NEC
	G035	Partial oesophagectomy and interposition of microvascularly attached colon
	G036	Partial oesophagectomy and interposition of colon NEC
	G038	Other specified partial excision of oesophagus
	G039	Unspecified partial excision of oesophagus
Stomach cancer (ICD10 C16)	G011	Oesophagogastrectomy and anastomosis of oesophagus to stomach
	G012	Oesophagogastrectomy and anastomosis of oesophagus to transposed jejunum
	G013	Oesophagogastrectomy and anastomosis of oesophagus to jejunum NEC
	G018	Other specified excision of oesophagus and stomach
	G019	Unspecified excision of oesophagus and stomach
	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
	G279	Unspecified total excision of stomach
	G281	Partial gastrectomy and anastomosis of stomach to duodenum
	G282	Partial gastrectomy and anastomosis of stomach to transposed jejunum
	G283	Partial gastrectomy and anastomosis of stomach to jejunum NEC
	G288	Other specified partial excision of stomach
	G289	Unspecified partial excision of stomach

Duodenal cancer (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 cancer (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 cancer (ICD10 C23)	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
	J188	Other specified excision of gall bladder
	J189	Unspecified excision of gall bladder
Biliary cancer (ICD10 C24)	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
Pancreatic cancer (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.