



Data quality report:

**Upper Gastrointestinal Cancer Site Specific
Clinical Reference Group (SSCRG)**

Victoria H Coupland
Julie Konfortion
Ruth H Jack
Karen M Linklater



Contents

1.	Introduction.....	3
2.	Methods.....	4
2.1	Data quality	4
3.	Results	7
3.1	Quality of the upper gastrointestinal cancer dataset, England, 2009.....	7
3.2	Death certificate only (DCO).....	8
3.3	Basis of diagnosis	10
3.4	Anatomical site	12
3.5	Morphology.....	14
3.6	Linked HES records	16
3.7	Ethnicity.....	17
3.8	T stage (pathological).....	20
3.9	N stage (pathological)	21
3.10	M stage (pathological)	22
3.11	TNM stage (pathological)	23
3.12	T stage (clinical)	24
3.13	N stage (clinical).....	25
3.14	M stage (clinical)	26
3.15	TNM stage (clinical).....	27
3.16	T stage (integrated)	28
3.17	N stage (integrated).....	29
3.18	M stage (integrated)	30
3.19	TNM stage (integrated).....	31
3.20	Nodes positive.....	32
3.21	Distant metastases.....	33
4.	Key findings.....	34
5.	Conclusions.....	35
	Appendix 1: List of ICD10 4 digit codes.....	36

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 liver, biliary tract, ampulla of Vater, duodenal, gallbladder and pancreatic cancers), (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 these 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 are being 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 2000 and 2009 focusing on the most recent diagnosis year (2009).

2. Methods

Data were extracted from the National Cancer Data Repository on all patients diagnosed with upper gastrointestinal cancers between 2000 and 2009. There were 238,028 tumours diagnosed in this ten-year period. Of these, 25,145 were diagnosed in 2009.

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 2009 by type of cancer, and the trends over time (2000-2009) 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. Four 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 and not known.

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 using ICD-O-2 (International Classification of Diseases for Oncology version 2) as known (valid morphology codes) and not known (morphology codes: 8000, 8001, 8010 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_yn" which records whether or not nodes that were found were 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 fields a value of X was recorded as valid not known. In the "nodes_postive_yn" and "mets" fields a value of Y or N were taken as valid known, and X was defined as valid not known.

3. Results

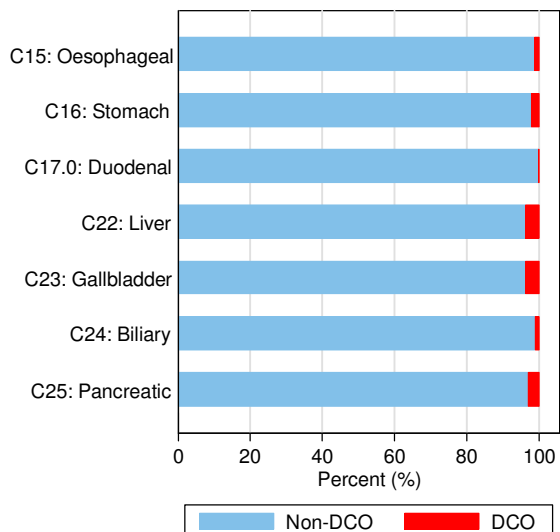
3.1 Quality of the upper gastrointestinal cancer dataset, England, 2009

	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)
Death certificate only (DCO)	6,711	6,153	393	3,538	574	751	7,025
DCO	82 (1.2)	135 (2.2)	1 (0.3)	135 (3.8)	21 (3.7)	8 (1.1)	217 (3.1)
Non-DCO	6,629 (98.8)	6,018 (97.8)	392 (99.7)	3,403 (96.2)	553 (96.3)	743 (98.9)	6,808 (96.9)
Basis of diagnosis (excluding DCO registrations)							
Microscopically verified	6,117 (92.3)	5,425 (90.1)	358 (91.3)	1,398 (41.1)	343 (62.0)	572 (77.0)	3,237 (47.5)
Clinically verified	477 (7.2)	561 (9.3)	33 (8.4)	1,953 (57.4)	204 (36.9)	168 (22.6)	3,483 (51.2)
Death certificate	20 (0.3)	27 (0.4)	1 (0.3)	31 (0.9)	4 (0.7)	1 (0.1)	62 (0.9)
Not known	15 (0.2)	5 (0.1)	0 (0.0)	21 (0.6)	2 (0.4)	2 (0.3)	26 (0.4)
Anatomical site (excluding DCO registrations)							
Known	3,702 (55.8)	3,308 (55.0)	-	-	-	649 (87.3)	3,373 (49.5)
Not known	2,927 (44.2)	2,710 (45.0)	-	-	-	94 (12.7)	3,435 (50.5)
Morphology (excluding DCO registrations)							
Known	5,963 (90.0)	5,334 (88.6)	340 (86.7)	2,685 (78.9)	350 (63.3)	627 (84.4)	3,103 (45.6)
Not known	666 (10.0)	684 (11.4)	52 (13.3)	718 (21.1)	203 (36.7)	116 (15.6)	3,705 (54.4)
Linked HES records (excluding DCO registrations)							
Linked	6,363 (96.0)	5,701 (94.7)	370 (94.4)	3,041 (89.4)	485 (87.7)	697 (93.8)	6,141 (90.2)
Not linked	266 (4.0)	317 (5.3)	22 (5.6)	362 (10.6)	68 (12.3)	46 (6.2)	667 (9.8)
Ethnicity (excluding DCO registrations)							
Known	6,181 (93.2)	5,523 (91.8)	364 (92.9)	2,948 (86.6)	468 (84.6)	679 (91.4)	5,884 (86.4)
Not known	448 (6.8)	495 (8.2)	28 (7.1)	455 (13.4)	85 (15.4)	64 (8.6)	924 (13.6)

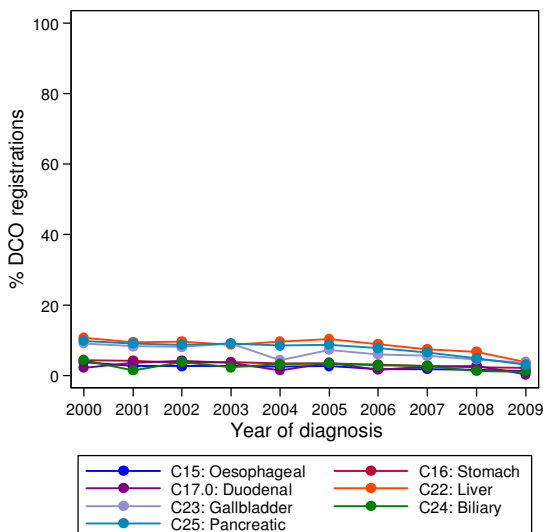
3.2 Death certificate only (DCO)

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

2009

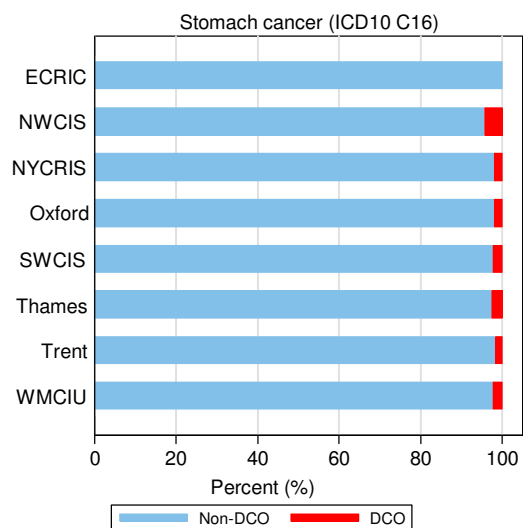
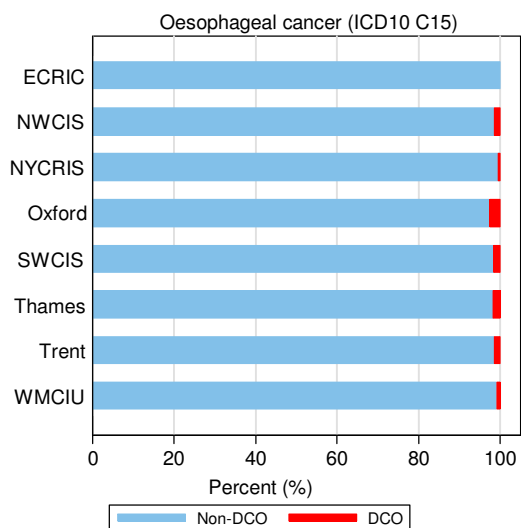


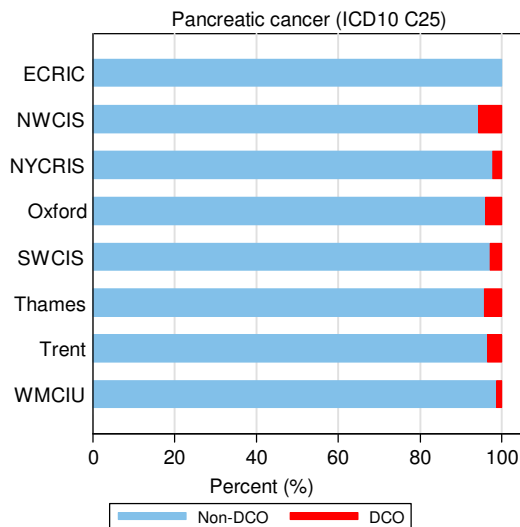
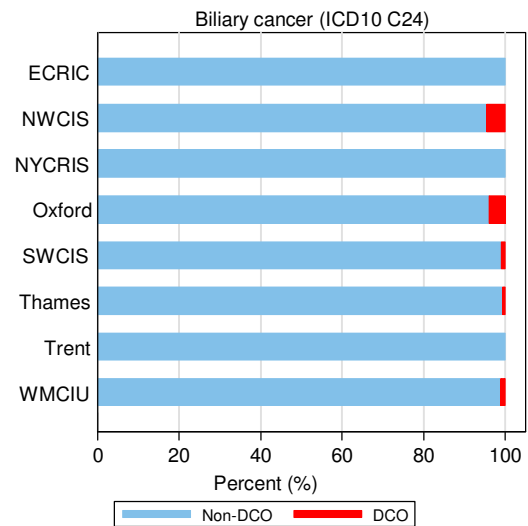
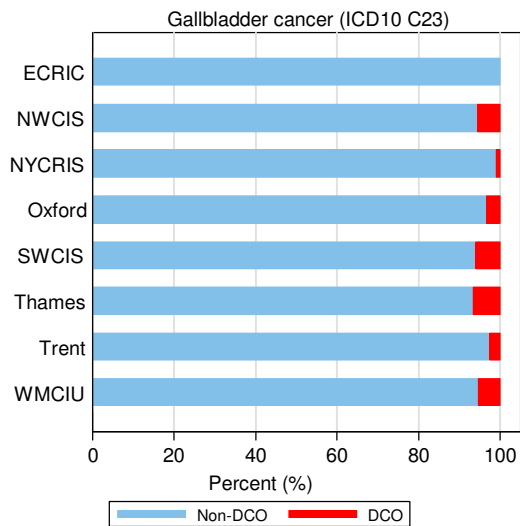
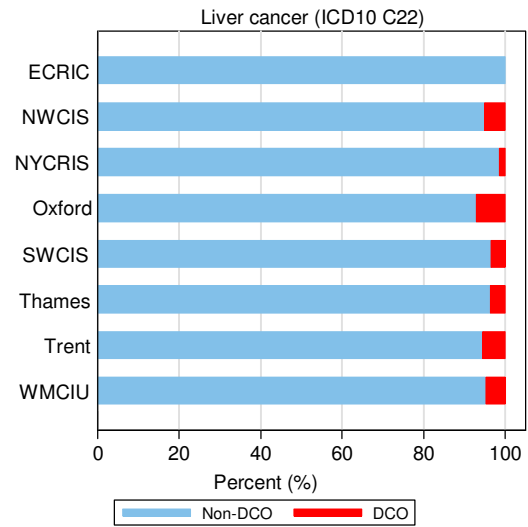
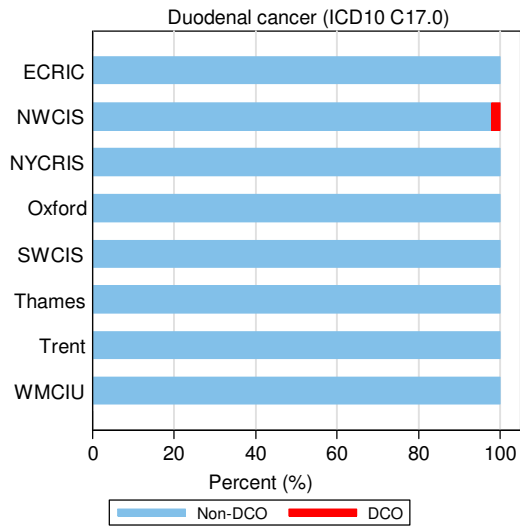
2000-2009



Less than 4% of all included cancer registrations were based on the death certificate only. The greatest proportion of DCO registrations was in liver cancer (3.8%). Between 2000 and 2009, the proportion of DCO registrations decreased for all cancer types.

2009



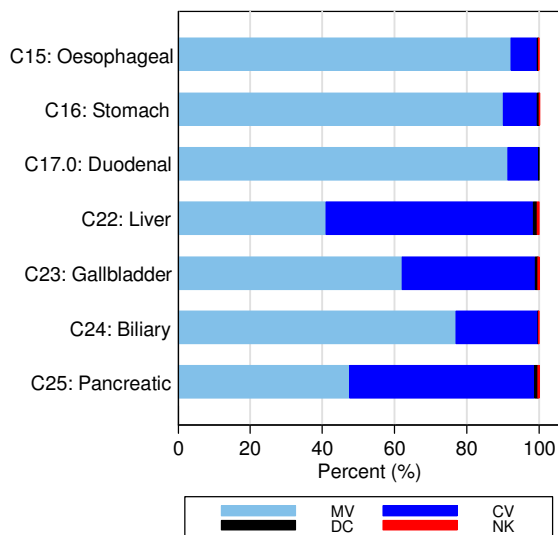


For most of the cancer types the proportion of DCO registrations was very low, with little variation between cancer registries

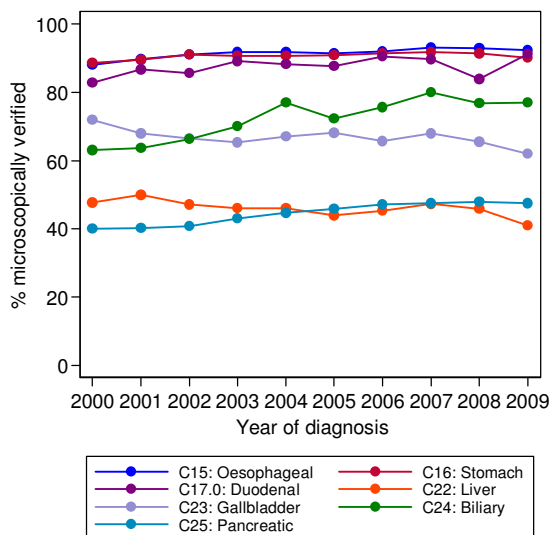
3.3 Basis of diagnosis

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

2009

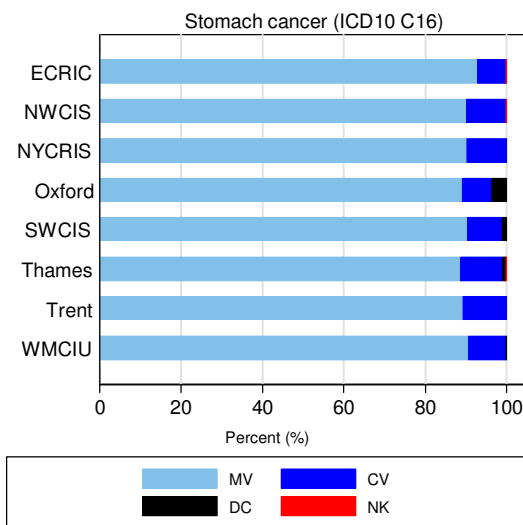
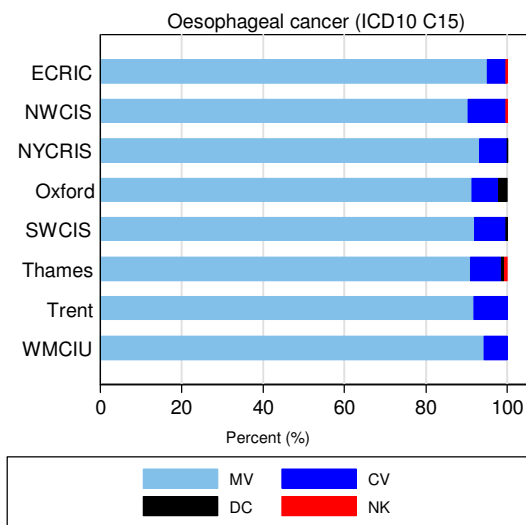


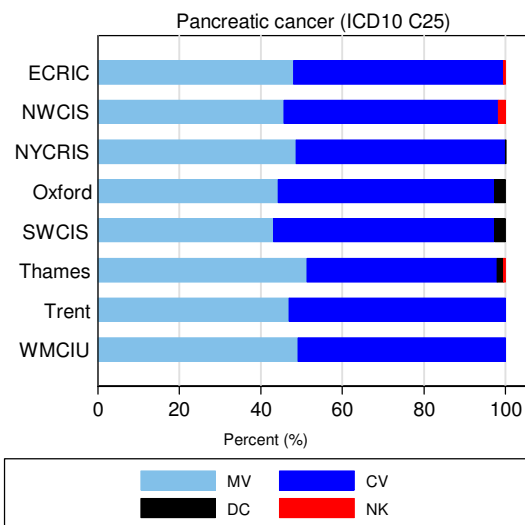
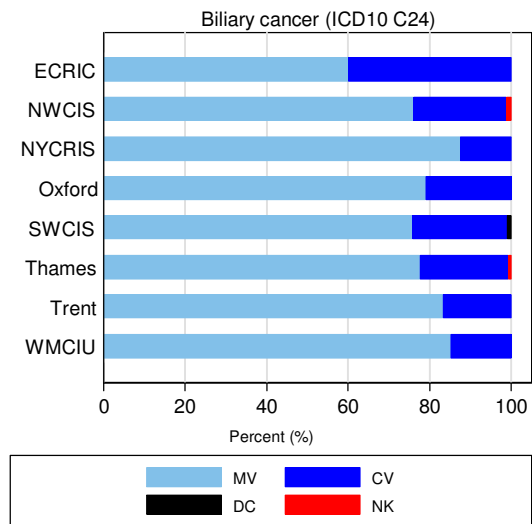
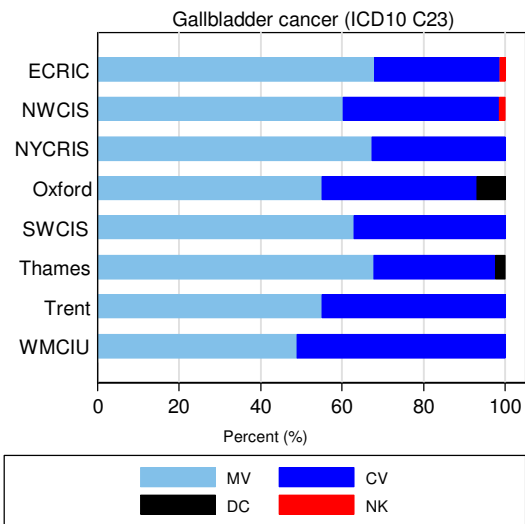
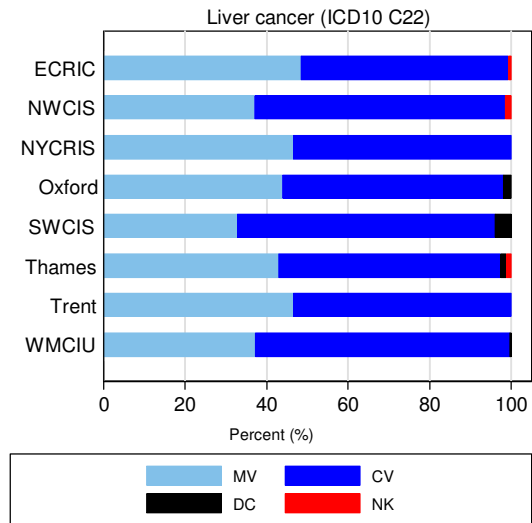
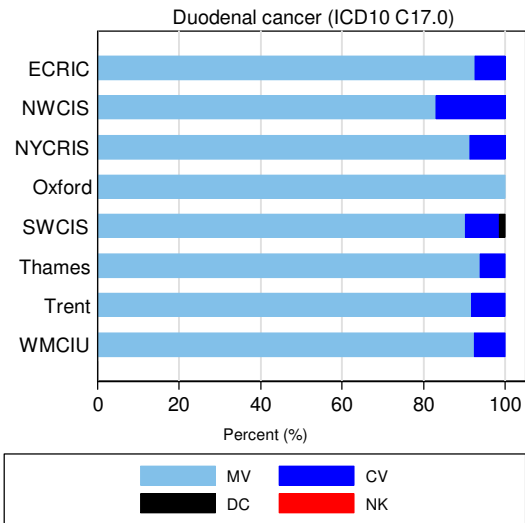
2000-2009



Over 90% of oesophageal and stomach cancers were microscopically verified. Liver (41.1%) and pancreatic (47.5%) cancers had the lowest proportions of microscopically verified cases. Over 50% of these cancers were clinically verified. Between 2000 and 2009 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.

2009



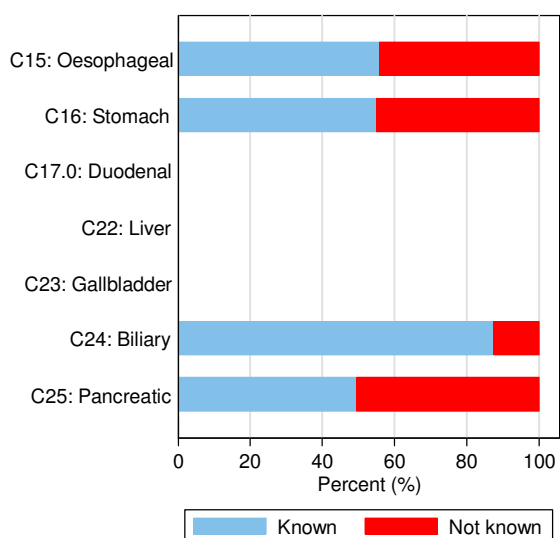


The lower proportion of microscopically verified hepatic, pancreatic and biliary cancers is probably due to these tumours being more inaccessible compared with the more accessible oesophageal and stomach cancers.

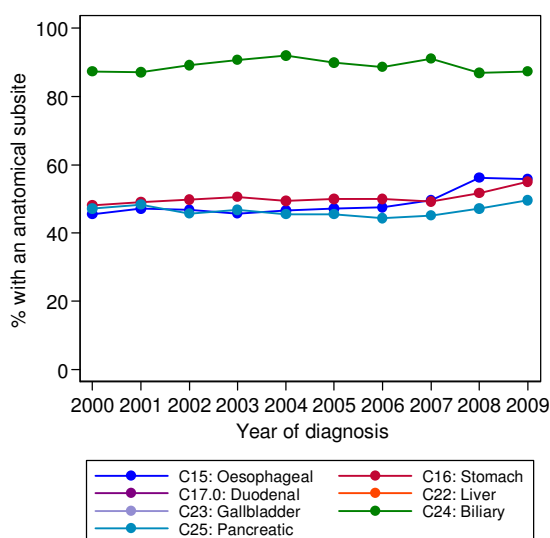
3.4 Anatomical site

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

2009

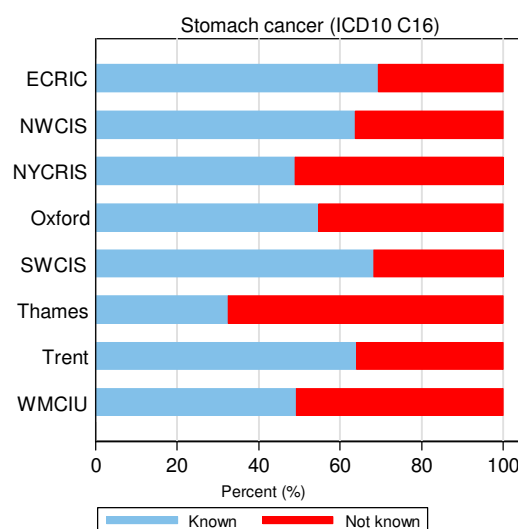
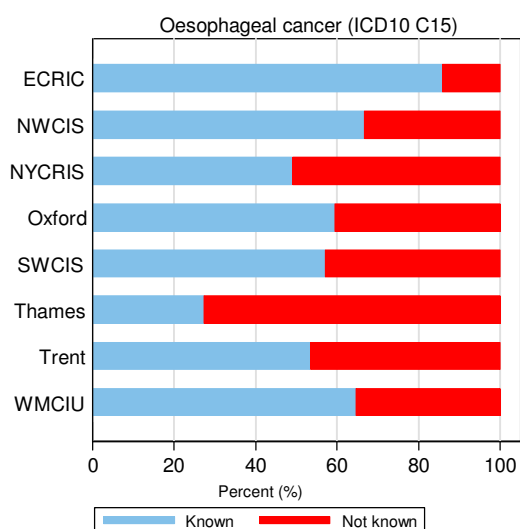


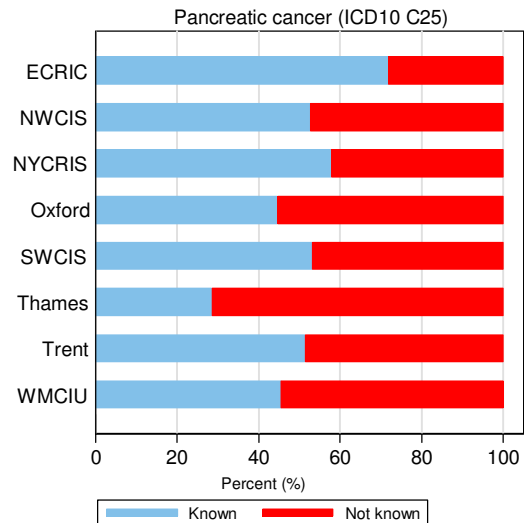
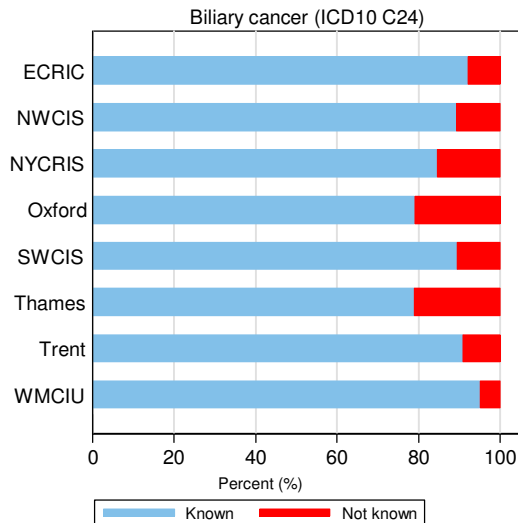
2000-2009



Around 50% of oesophageal, stomach and pancreatic and 87% of biliary cancer registrations had a known anatomical subsite. The proportion of oesophageal cancer registrations with a known anatomical subsite increased between 2000 and 2009. A relatively stable trend was found for the other cancer types.

2009



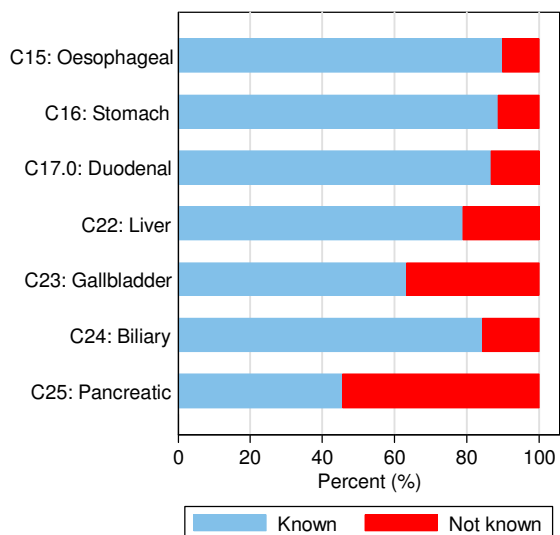


Duodenal, liver and gallbladder cancer were not included in this section. Duodenal cancer is defined by the ICD10 4 digit code of C17.0 (see Appendix 1). Those with an unspecified anatomical location in the C17 (malignant neoplasm of the small intestine) group are defined as C17.8 (overlapping lesion 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.

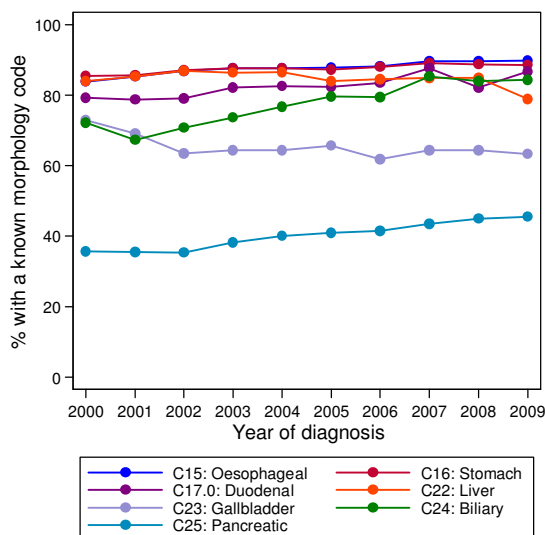
3.5 Morphology

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.

2009

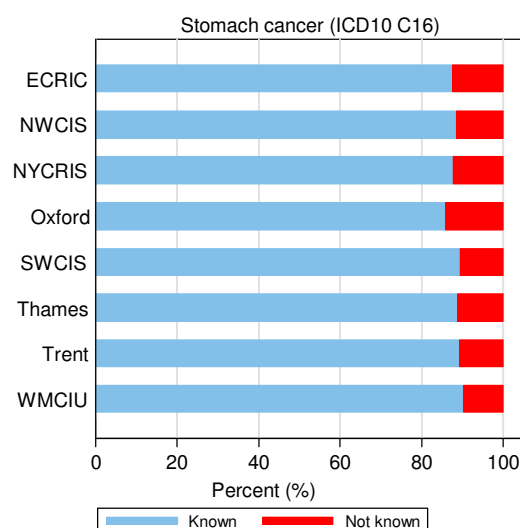
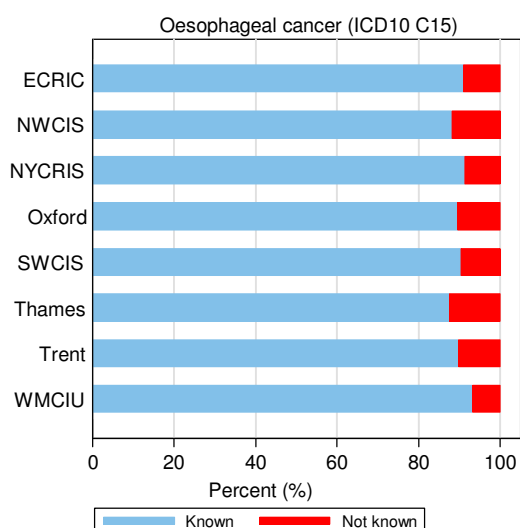


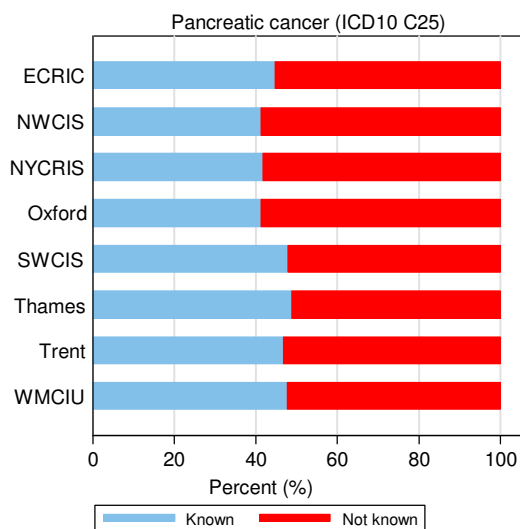
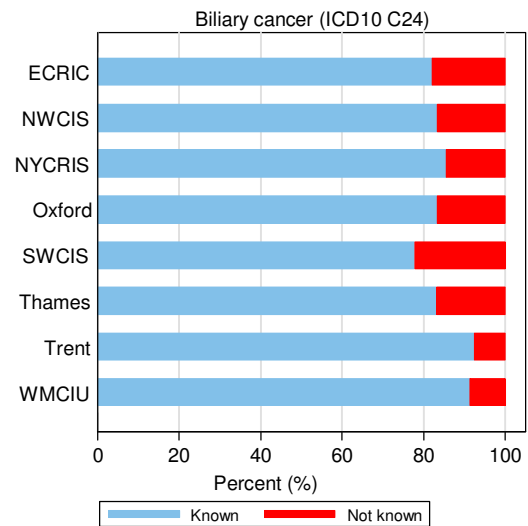
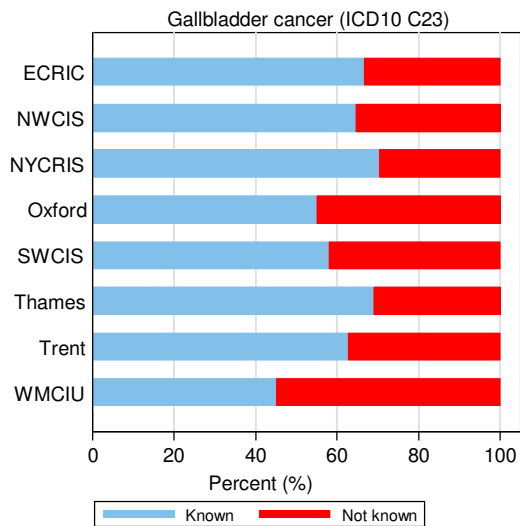
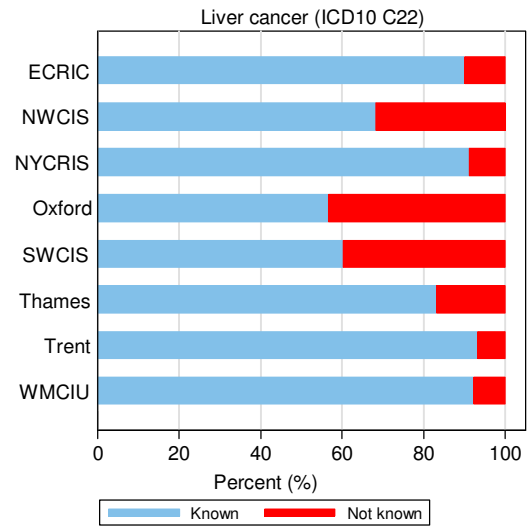
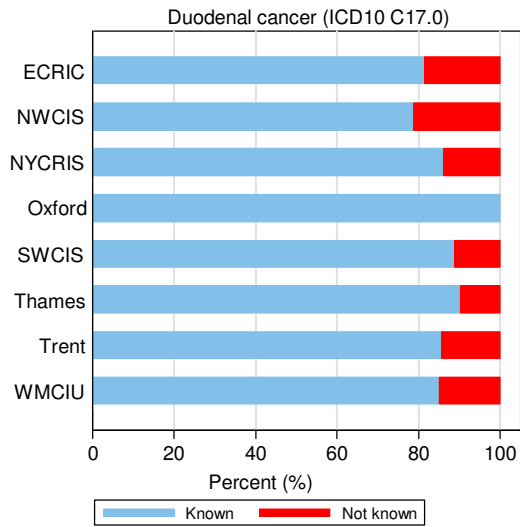
2000-2009



The highest proportion of registrations with known morphology information was found in oesophageal and stomach cancer (over 88%). Pancreatic (45.6%) and gallbladder (63.3%) cancer had the lowest proportion with a known morphology. Between 2000 and 2009 there was a relatively stable trend in the proportions of registrations with a known morphology for most of the cancer types, although biliary cancer increased from 72.24% to 84.36% and pancreatic cancer from 35.7% to 45.6%.

2009



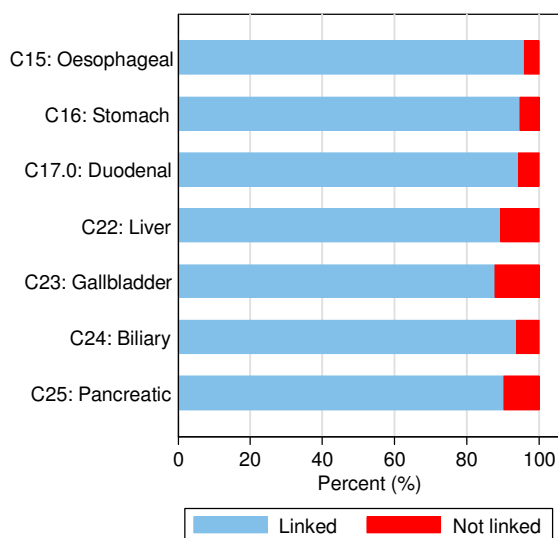


There was some variation in the proportion of registrations with a known morphology between cancer registries.

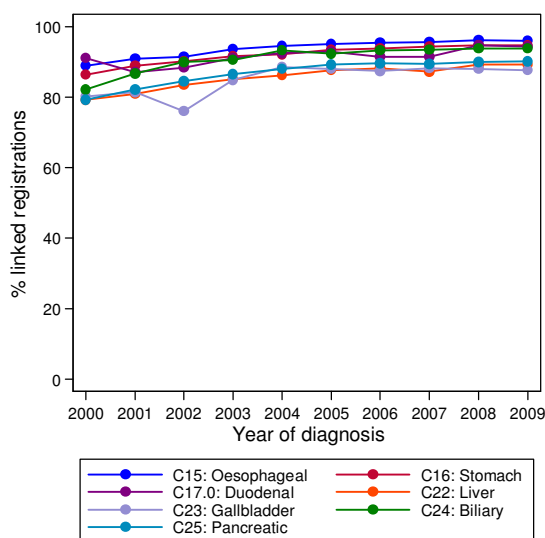
3.6 Linked HES records

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.

2009

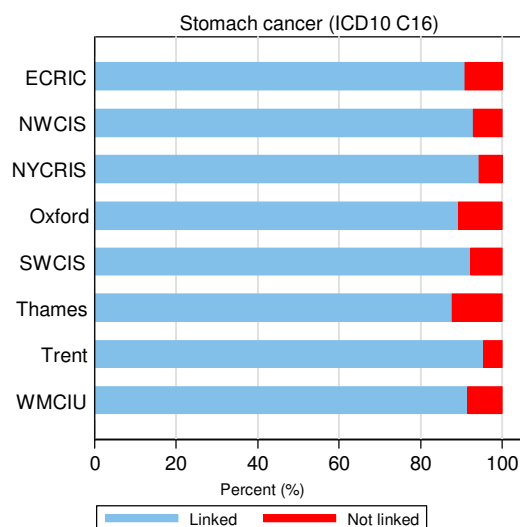
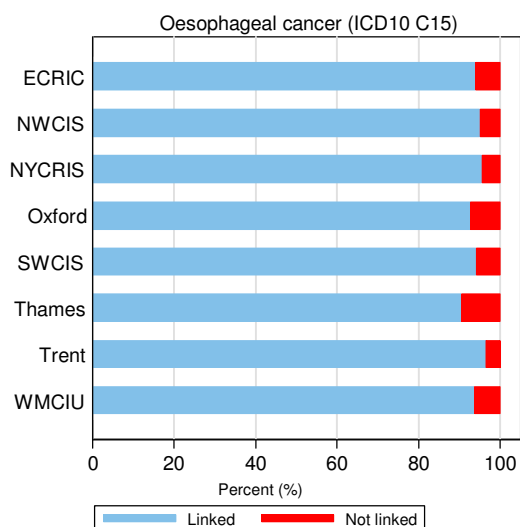


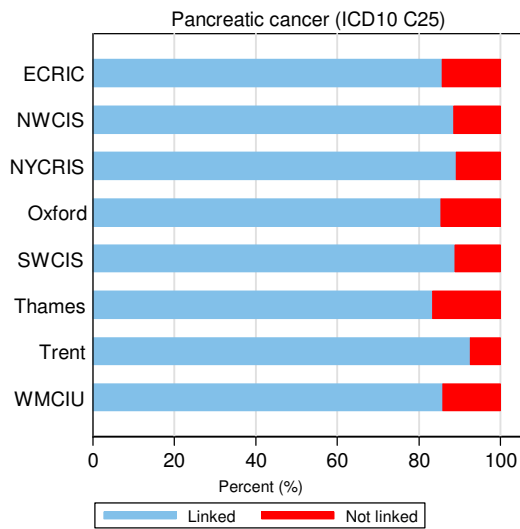
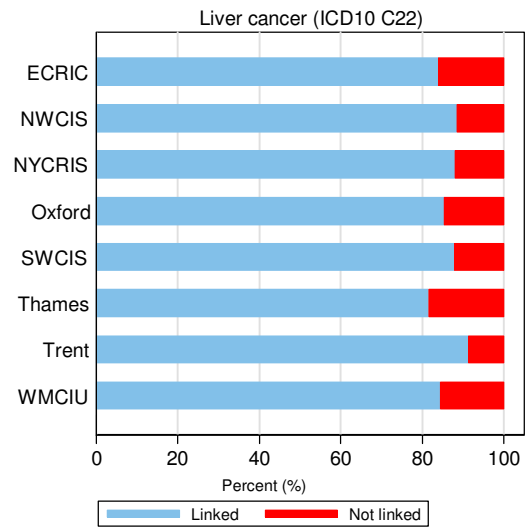
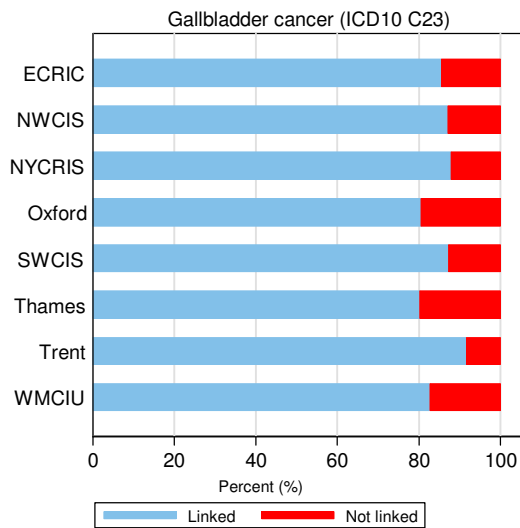
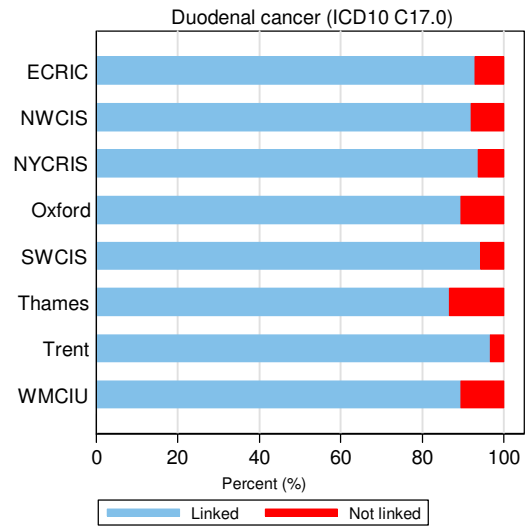
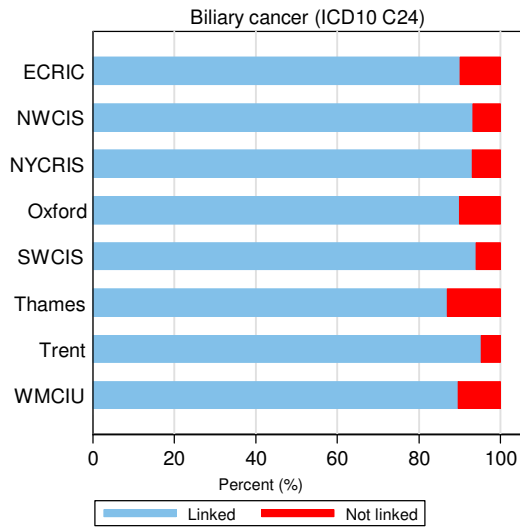
2000-2009



Over 90% of oesophageal, stomach, duodenal, biliary and pancreatic cancer registrations had a linked HES record in 2009. Gallbladder (87.7%) and liver (89.4%) cancer had a lower proportion with a matched HES record. Between 2000 and 2009 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 2000.

2009



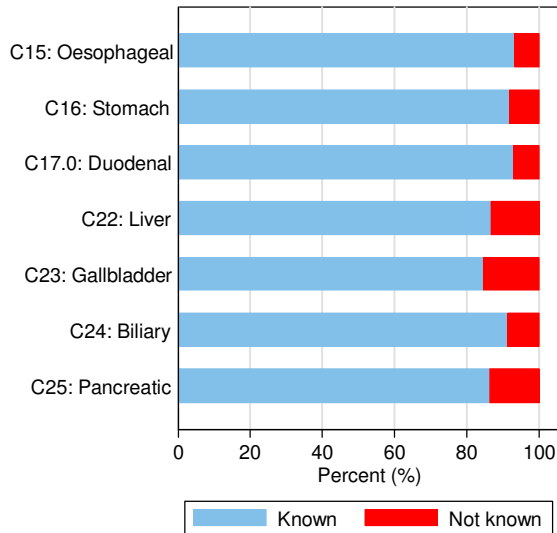


There was some variation in the proportion of registrations with a linked HES record between cancer registries.

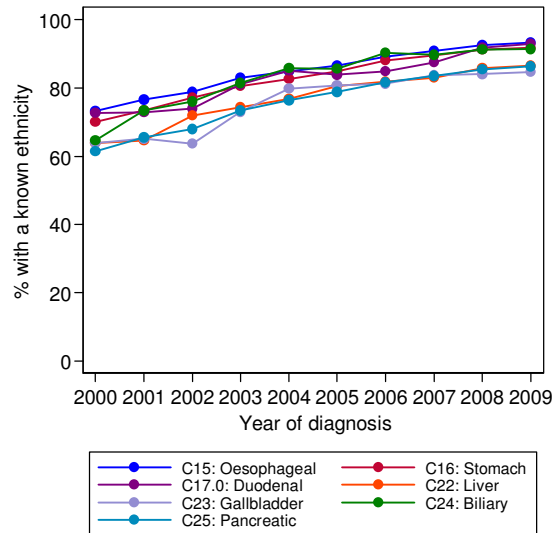
3.7 Ethnicity

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.

2009

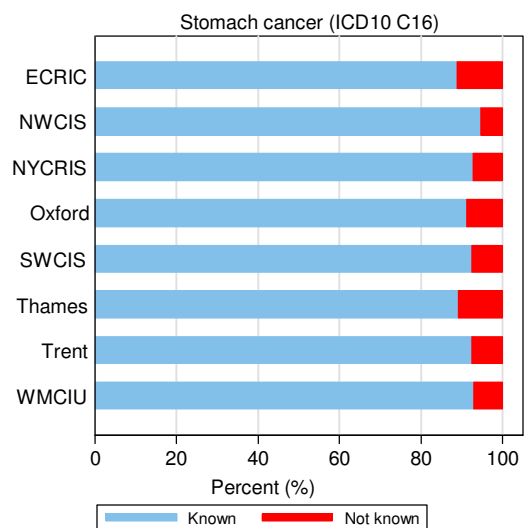
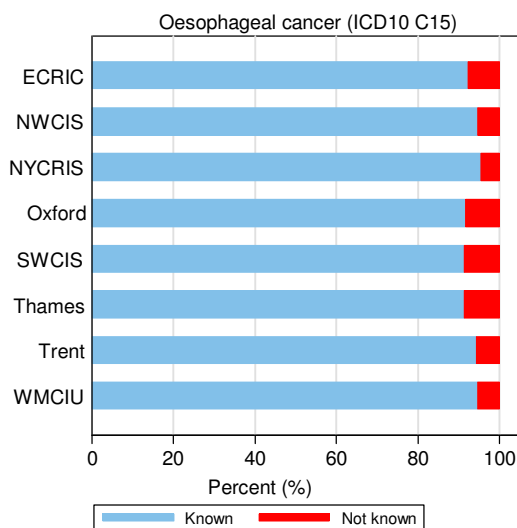


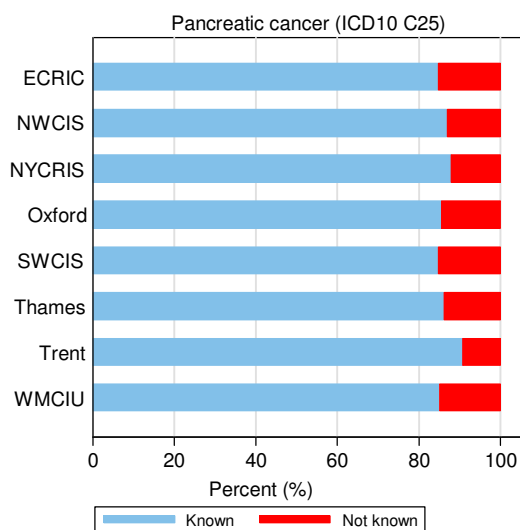
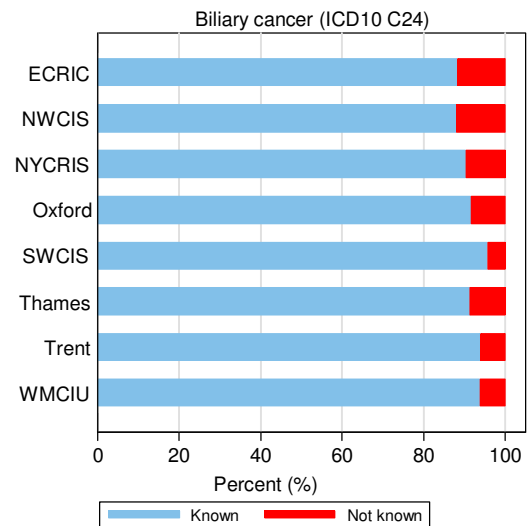
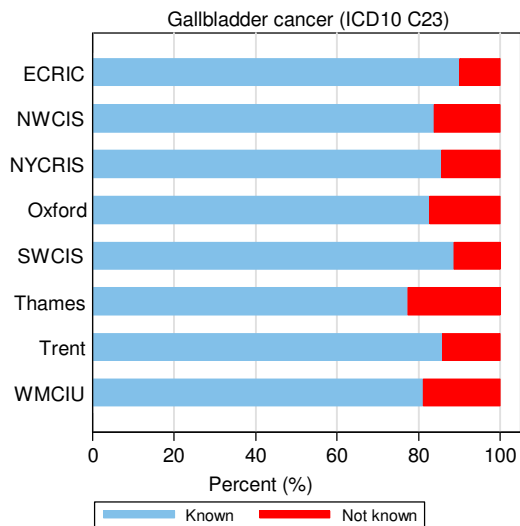
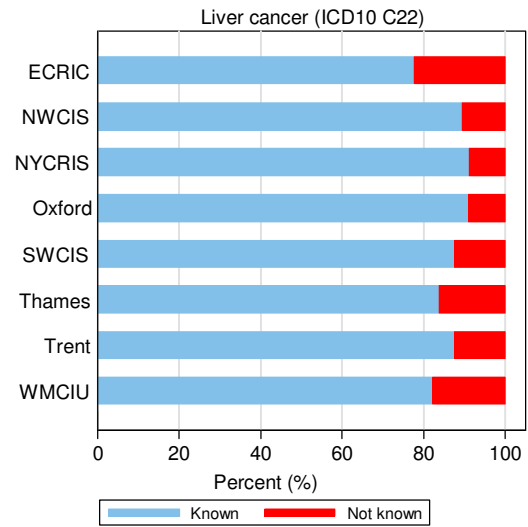
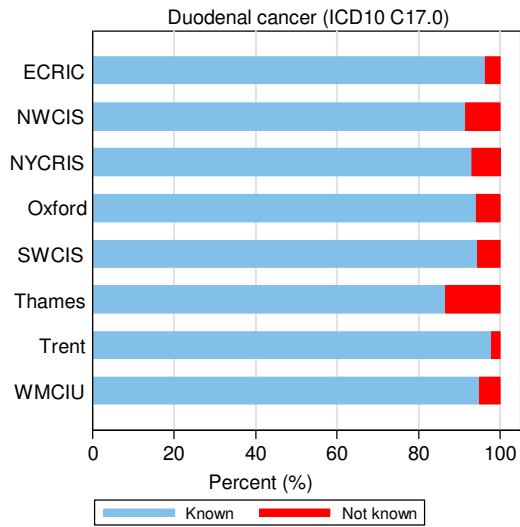
2000-2009



Overall, a high proportion of registrations had a known ethnicity. Liver (86.6%), pancreatic (86.4%) and gallbladder (84.6%) cancer had the lowest proportion of registrations with a known ethnicity. Between 2000 and 2009 the proportion of registrations with a known ethnicity increased for all cancer types.

2009





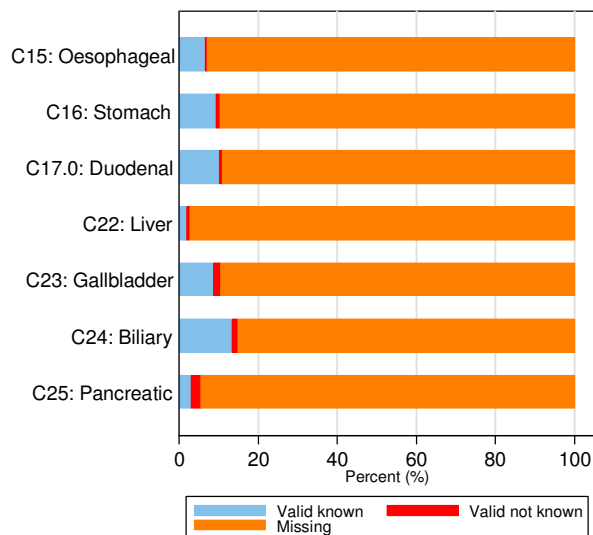
There was some variation in the proportion of registrations with a known ethnicity between cancer registries.

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

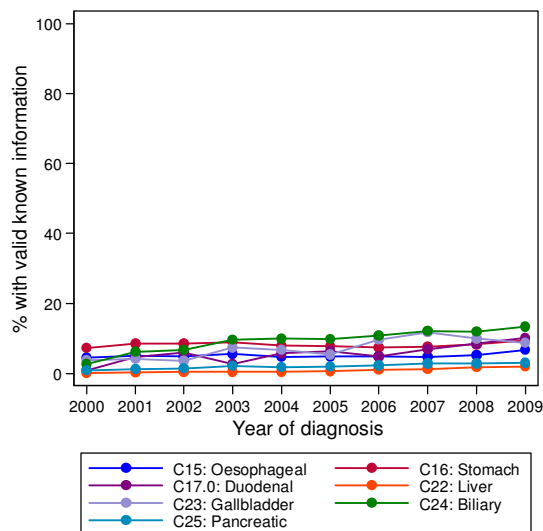
3.8 T stage (pathological)

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.

2009

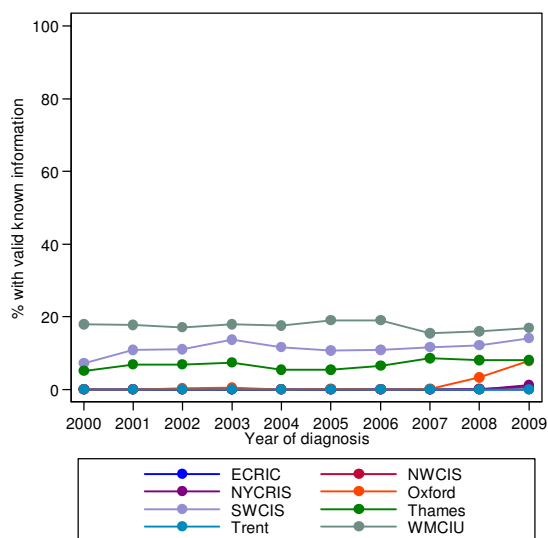


2000-2009



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 (13.5%). Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known T stage for most of the cancer types.

2000-2009



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

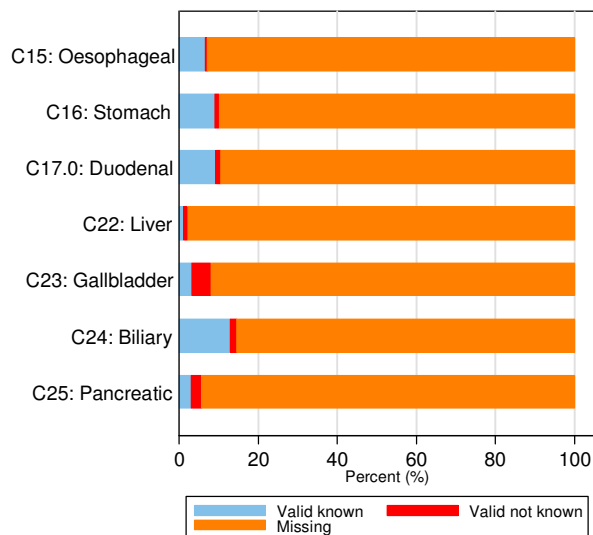
Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with a valid known T (pathological) stage across most cancer registries.

Since 2007, the proportion of registrations with a valid known T (pathological) stage increased in Oxford.

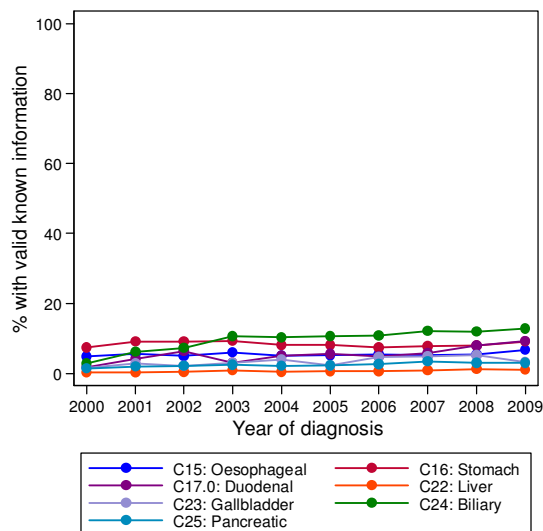
3.9 N stage (pathological)

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.

2009

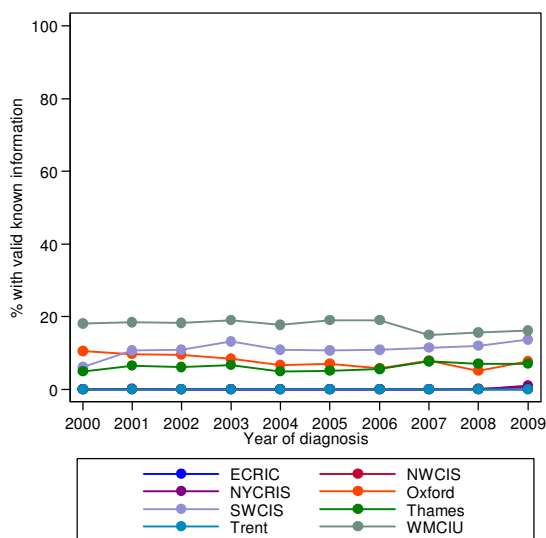


2000-2009



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 (12.9%). Between 2000 and 2009 there was an increase in the proportion of registrations with a valid known N stage for most cancer types.

2000-2009



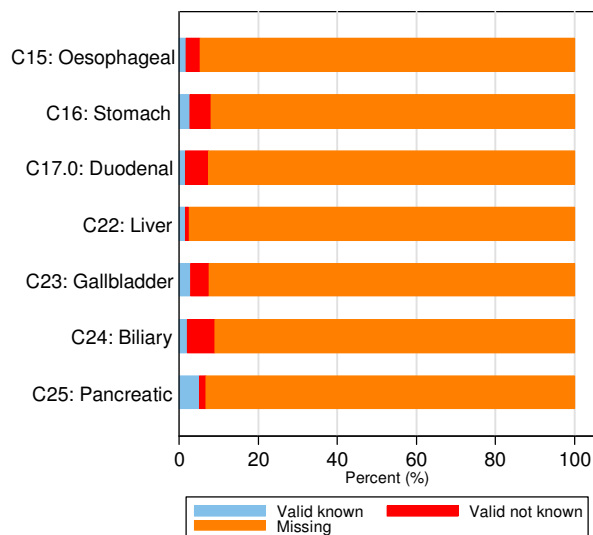
Not all cancer registries submitted their staging information in the N (pathological) stage field.

Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with a valid known N (pathological) stage across most cancer registries.

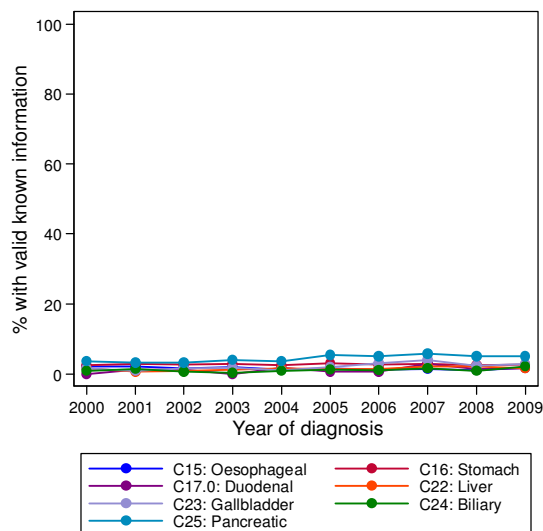
3.10 M stage (pathological)

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.

2009

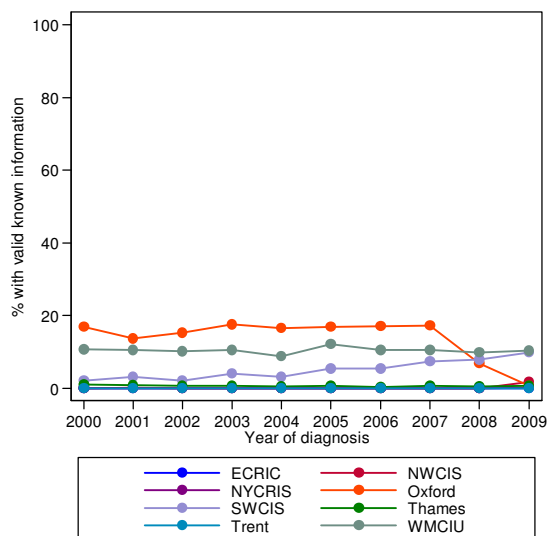


2000-2009



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 (5.1%). Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with a valid known M stage for some cancer types.

2000-2009



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

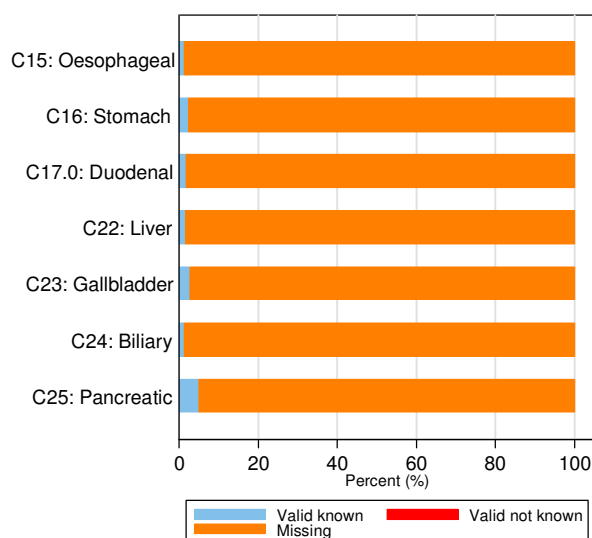
Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with a valid known M (pathological) stage across most cancer registries.

Since 2007, the proportion of registrations with a valid known M (pathological) stage decreased in Oxford.

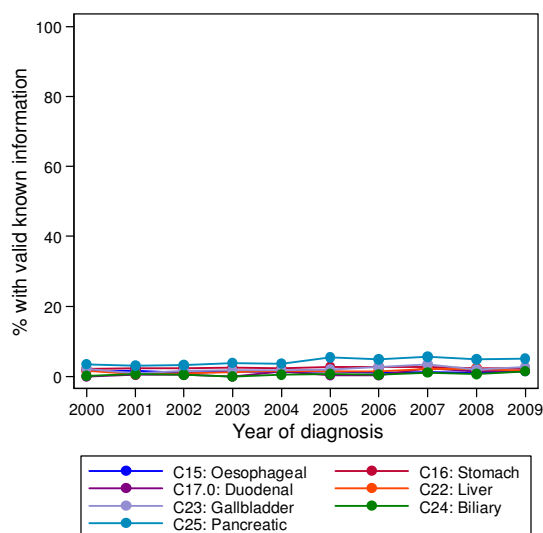
3.11 TNM stage (pathological)

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.

2009

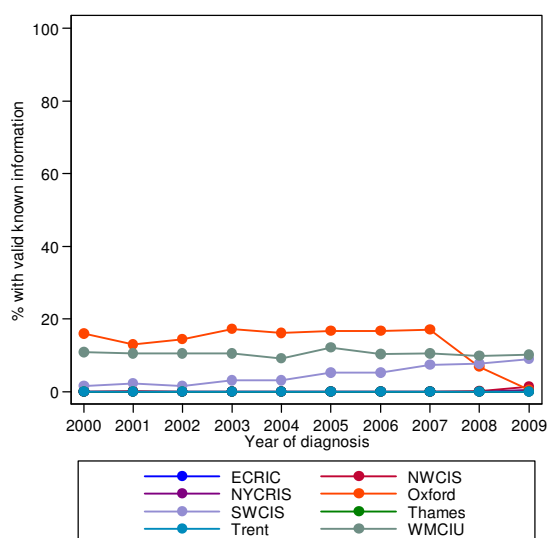


2000-2009



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 (5.0%). Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with a valid known TNM stage for all cancer types.

2000-2009



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

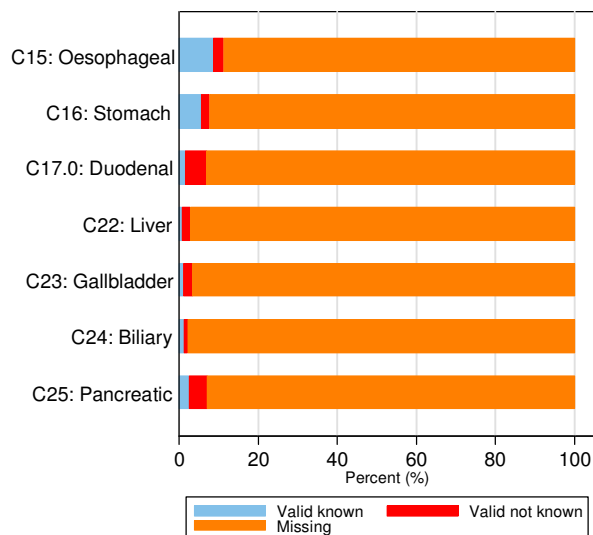
Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with a valid known TNM (pathological) stage across most cancer registries.

Since 2007, the proportion of registrations with a valid known TNM (pathological) stage decreased in Oxford.

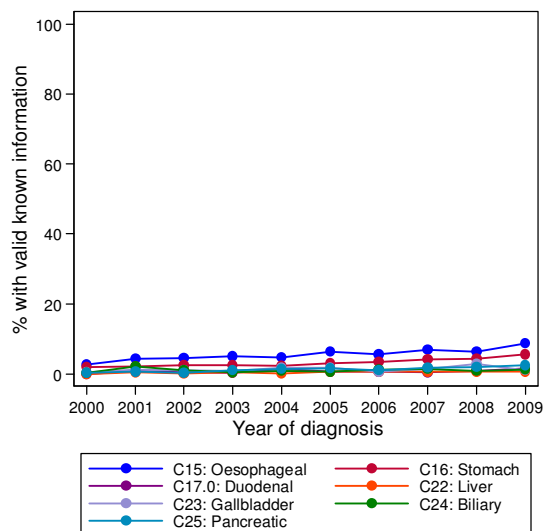
3.12 T stage (clinical)

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.

2009

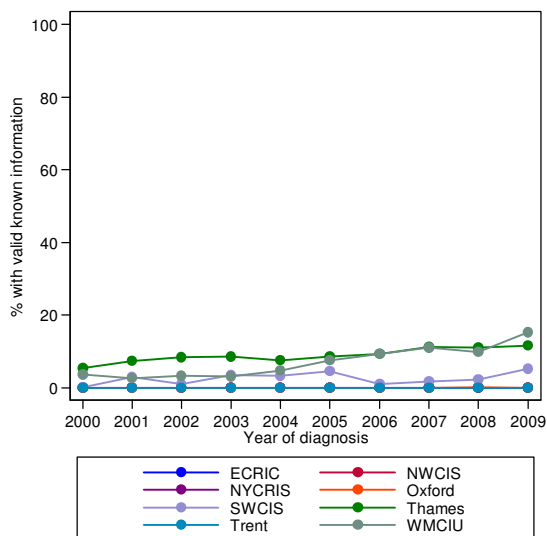


2000-2009



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 (8.7%). Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known T stage for some cancer types.

2000-2009



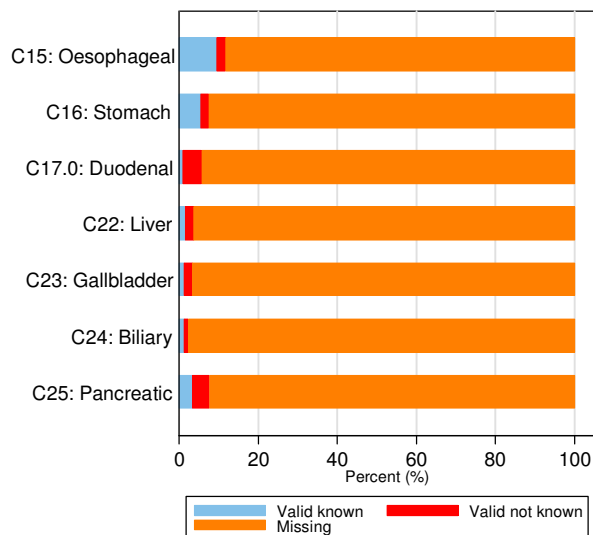
Not all cancer registries submitted their staging information in the T (clinical) stage field.

Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known T (clinical) stage in Thames and WMCIU.

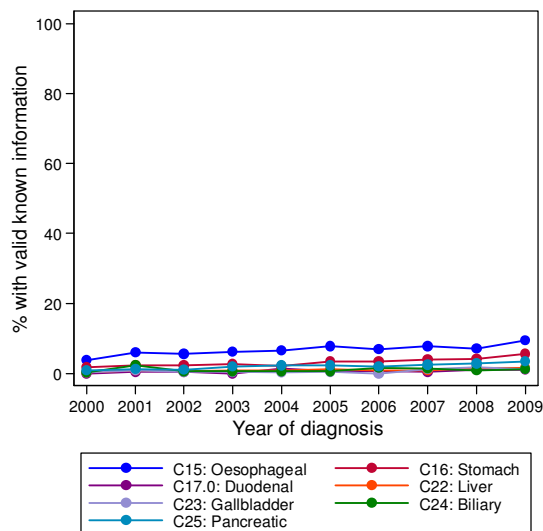
3.13 N stage (clinical)

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.

2009

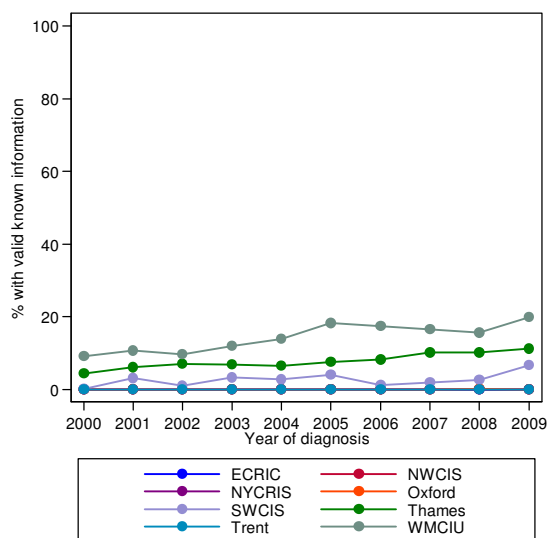


2000-2009



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 (9.5%). Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known N stage for some of the cancer types.

2000-2009



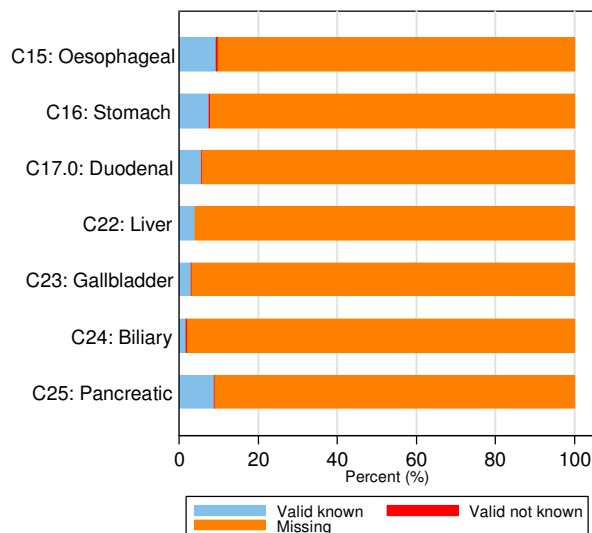
Not all cancer registries submitted their staging information in the N (clinical) stage field.

Between 2000 and 2009 there was an increase in the proportion of registrations with a valid known N (clinical) stage in SWCIS, Thames and WMCIU.

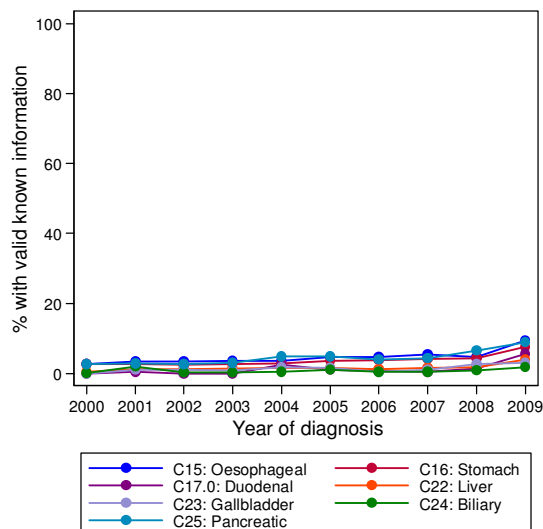
3.14 M stage (clinical)

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.

2009

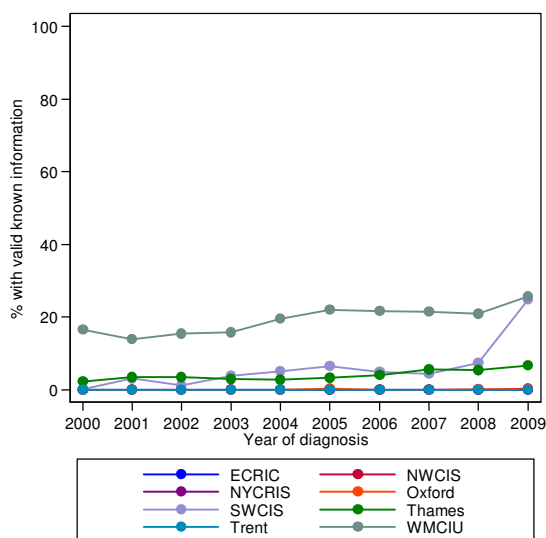


2000-2009



Overall, there were low proportions of M (clinical) stage recorded across each cancer type. Oesophageal and pancreatic cancer had the highest proportion of registrations with a valid known M stage. Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known M stage for most cancer types.

2000-2009



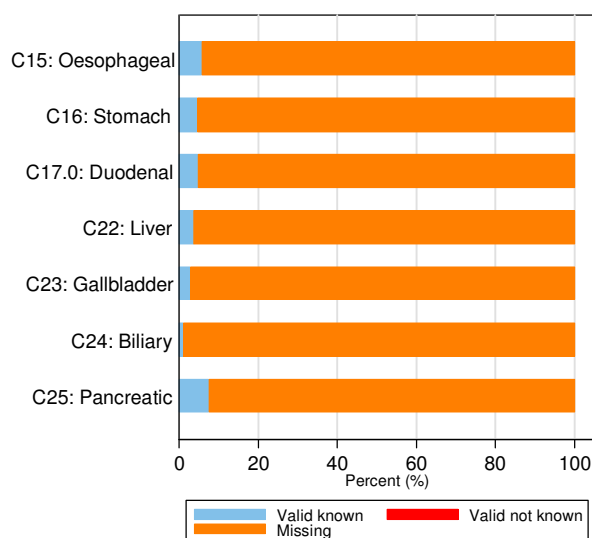
Not all cancer registries submitted their staging information in the M (clinical) stage field.

Between 2008 and 2009 there was a notable increase in the proportion of registrations with a valid known M (clinical) stage in SWCIS.

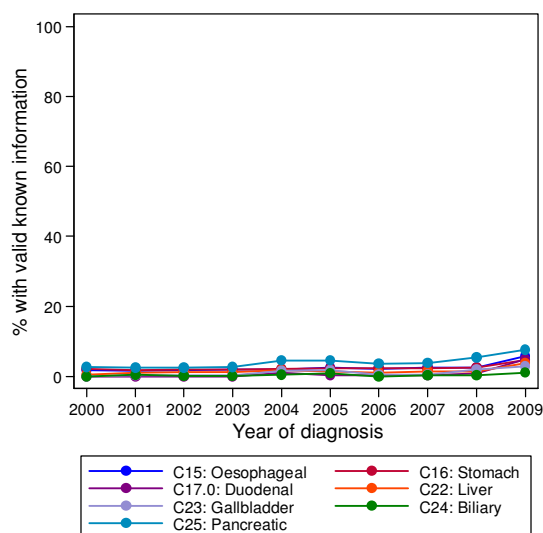
3.15 TNM stage (clinical)

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.

2009

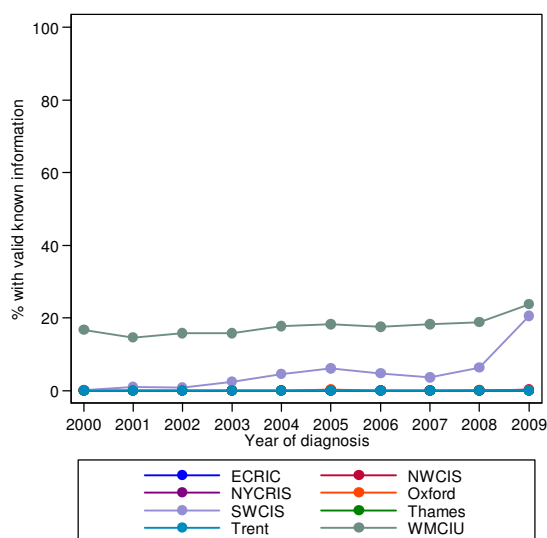


2000-2009



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 (7.6%). Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known TNM stage for some cancer types.

2000-2009



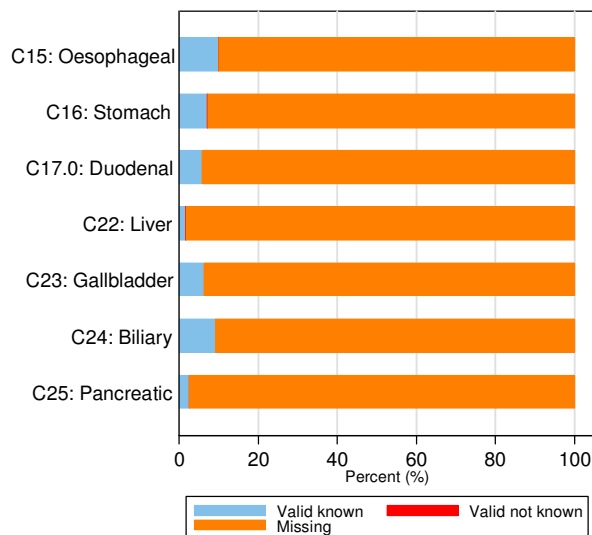
Not all cancer registries submitted their staging information in the TNM (clinical) stage field.

Between 2008 and 2009 there was a notable increase in the proportion of registrations with a valid known TNM (clinical) stage in SWCIS.

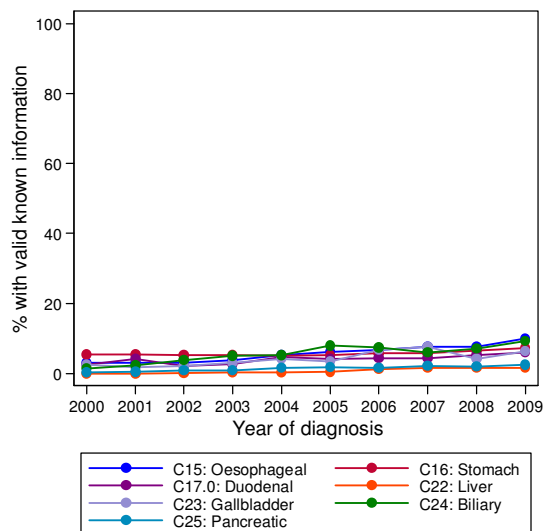
3.16 T stage (integrated)

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.

2009

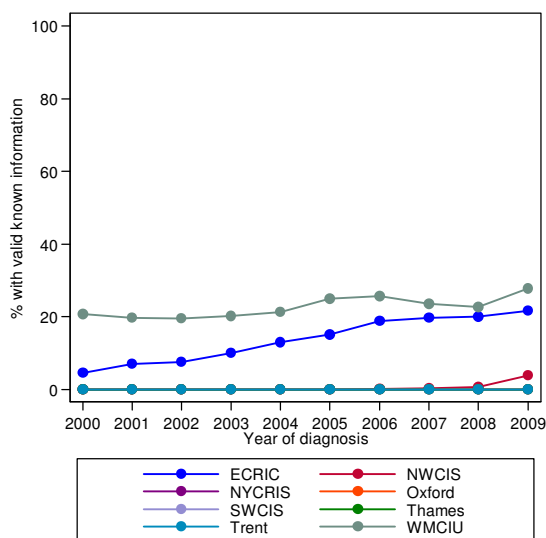


2000-2009



Overall, there were low proportions of T (integrated) stage recorded across each cancer type. Oesophageal and biliary cancer had the highest proportions of registrations with a valid known T stage above 7%. Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known T stage for some cancer types.

2000-2009



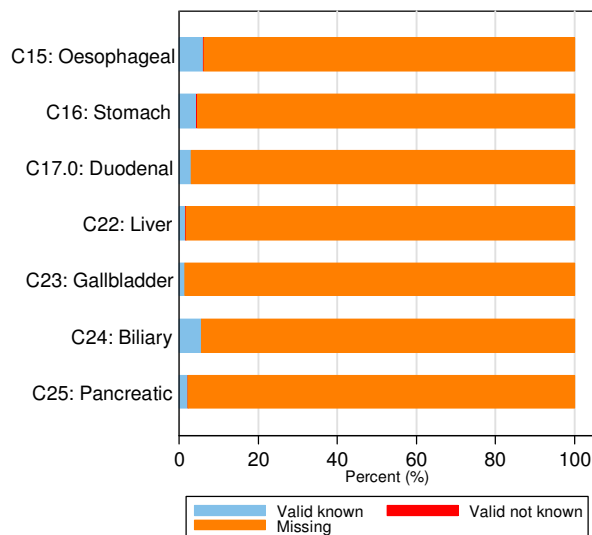
Three cancer registries (ECRIC, WMCIU and NWCIS) have submitted their staging information using the T (integrated) stage field.

Since 2000, there was an increase in the proportion of registrations with a valid known T (integrated) stage in ECRIC.

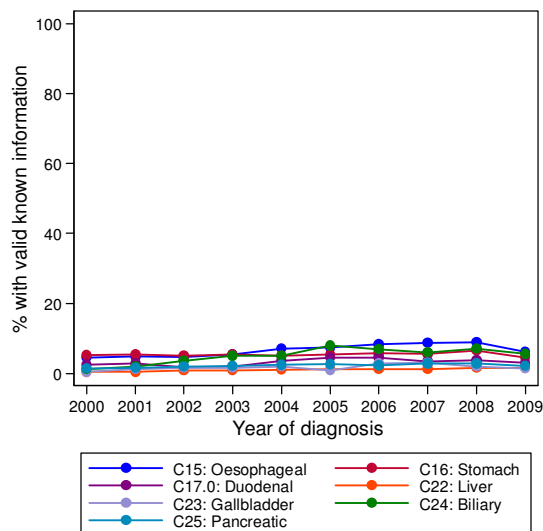
3.17 N stage (integrated)

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.

2009

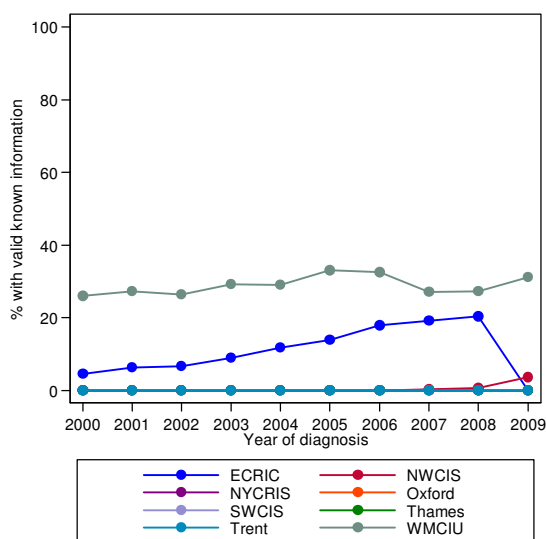


2000-2009



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 (6.1%). Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known N stage for some cancer types.

2000-2009



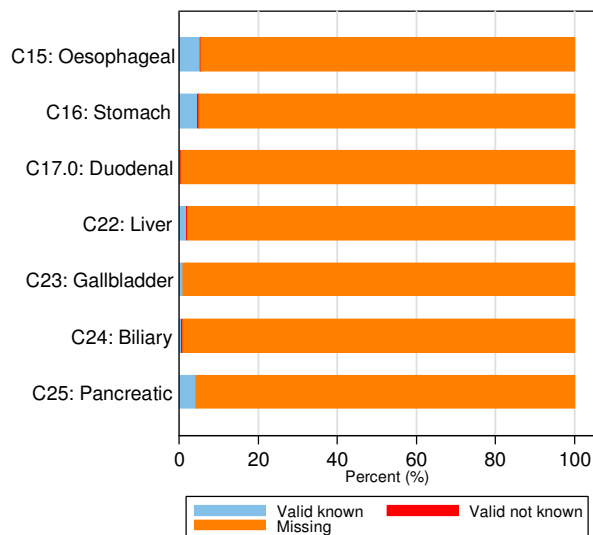
Three cancer registries (ECRIC, WMCIU and NWCIS) have submitted their staging information using the N (integrated) stage field.

In 2009, ECRIC had no information in the N (integrated) stage field.

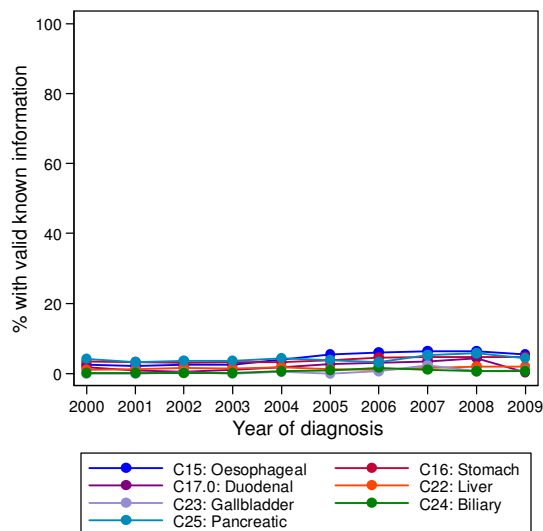
3.18 M stage (integrated)

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.

2009

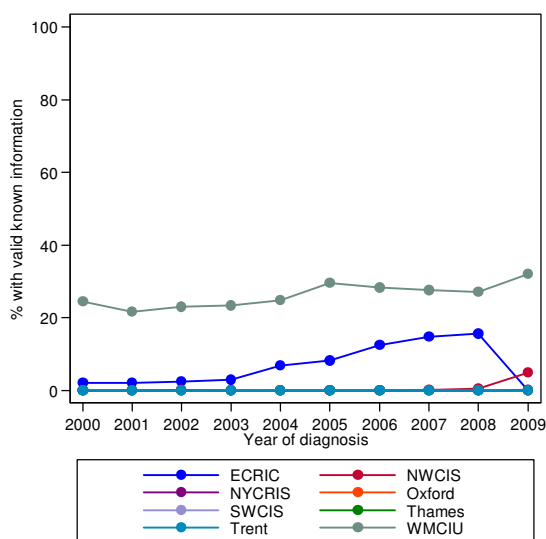


2000-2009



Overall, there were low proportions of M (integrated) stage recorded across each cancer type. Oesophageal, stomach and pancreatic cancers had the highest proportion of registrations with a valid known M stage above 4%. Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known M stage for some cancer types.

2000-2009



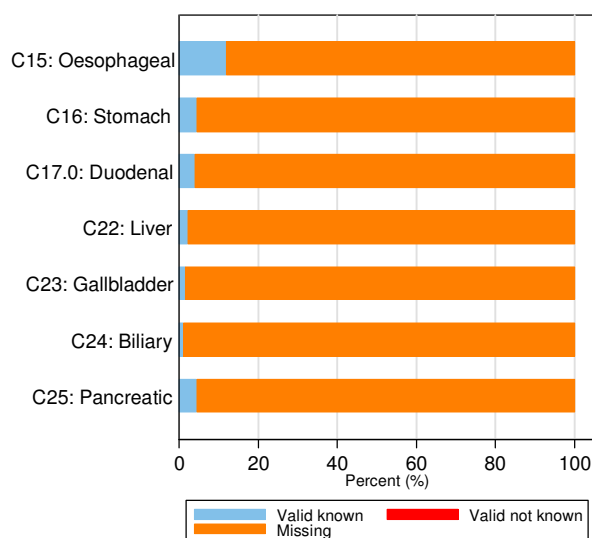
Three cancer registries (ECRIC, WMCIU and NWCIS) have submitted their staging information using the M (integrated) stage field.

In 2009, ECRIC had no information in the M (integrated) stage field.

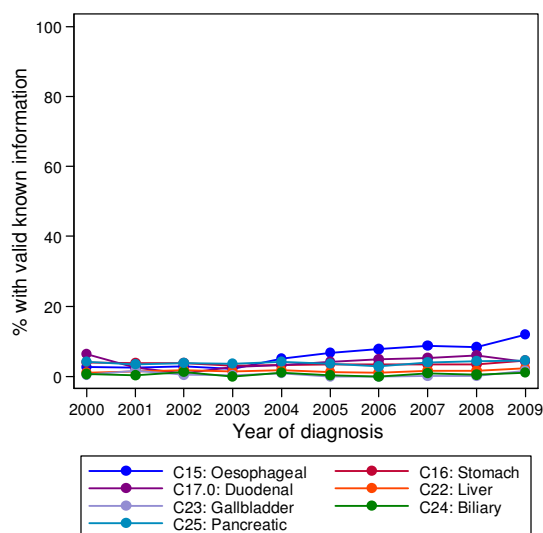
3.19 TNM stage (integrated)

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.

2009

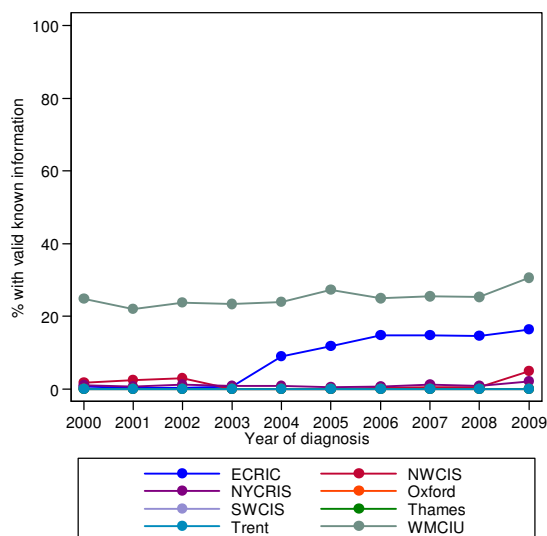


2000-2009



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 (11.9%). Between 2000 and 2009 there was a slight increase in the proportion of registrations with a valid known TNM stage for some cancer types.

2000-2009



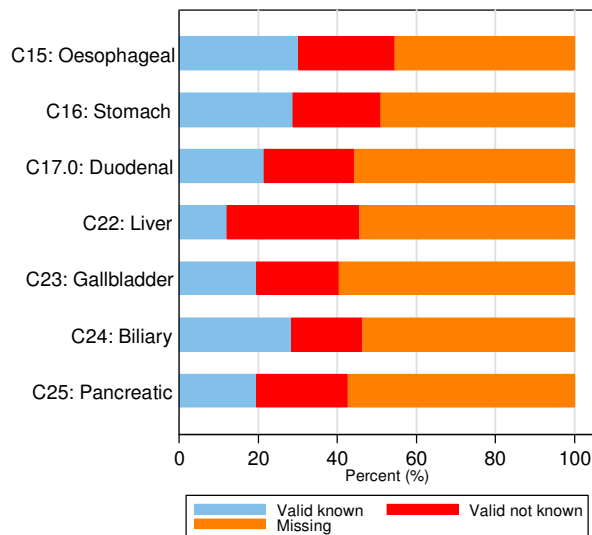
Four cancer registries (ECRIC, WMCIU, NYCRIS and NWCIS) have submitted their staging information using the TNM (integrated) stage.

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

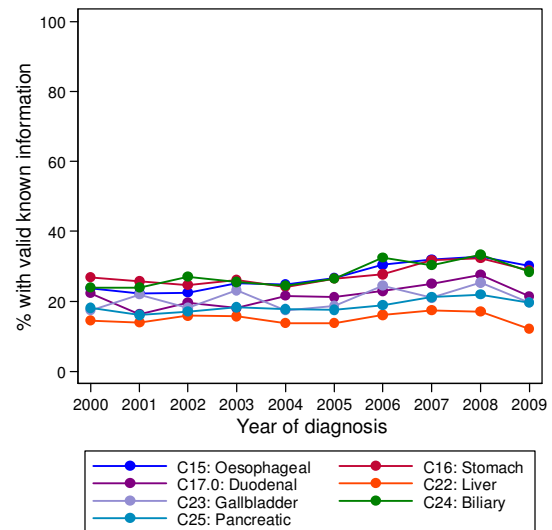
3.20 Nodes positive

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.

2009

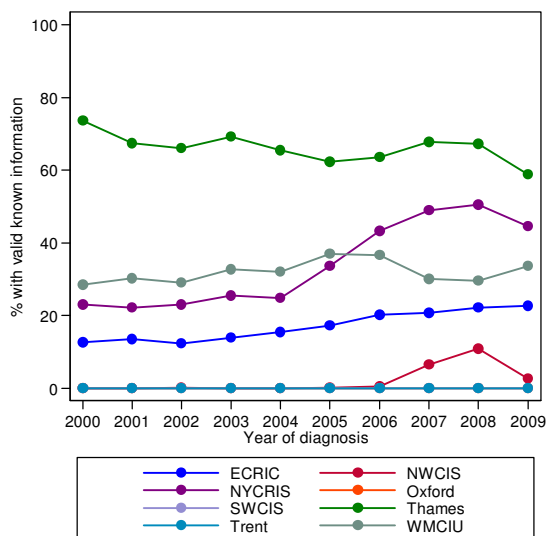


2000-2009



Overall, there were low proportions of nodes positive information recorded across each cancer type. Oesophageal cancer had the highest proportion of registrations with valid known nodes positive information (30.1%). Between 2000 and 2009 there was a relatively stable trend in the proportion of registrations with valid known nodes positive information for all cancer types.

2000-2009

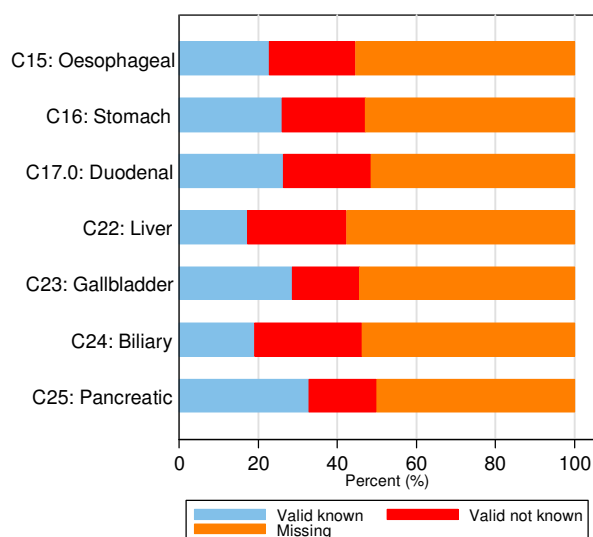


In general, there was an increase in the proportion of registrations with valid known nodes positive information in ECRIC and NYCRIS, and a decrease in Thames between 2000 and 2009.

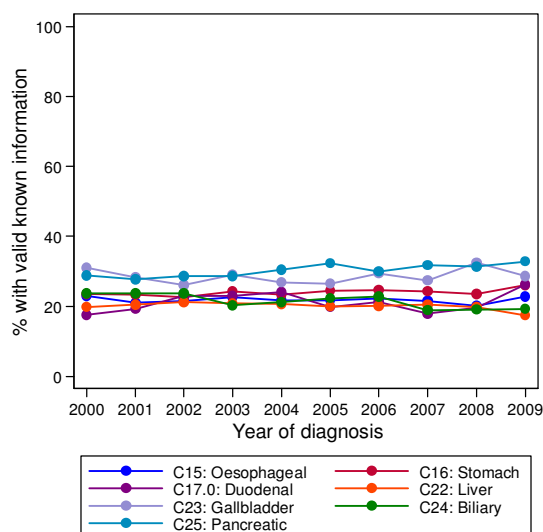
3.21 Distant metastases

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.

2009

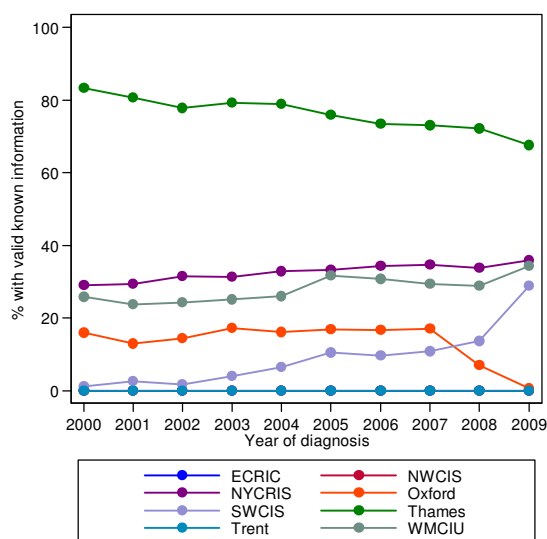


2000-2009



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 (32.9%). Between 40% and 45% of registrations were either valid known and valid not known across all cancer types. Between 2000 and 2009 there was a stable trend in the proportion of registrations with valid known metastases information.

2000-2009



Thames had the highest proportion of registrations with valid known metastases information, although this proportion declined between 2000 and 2009.

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

4. Key findings

- In 2009, the proportion of death certificate only registrations ranged from 0.3% (duodenal cancer) to 3.8% (liver cancer). Primary liver and stomach cancer had the highest proportions of DCO registrations. The proportion of DCO registrations decreased over time (2000-2009).
- The proportions of microscopically verified cases ranged from 41.1% (liver cancer) to 92.3% (oesophageal cancer) in 2009. Smaller proportions of registrations were microscopically verified in primary liver (41.1%) and pancreatic (47.5%) cancers compared with oesophageal and stomach cancers. Over half of liver (57.4%) and pancreatic (51.2%) cancers were only clinically verified.
- Around half of oesophageal (55.8%), stomach (55.0%) and pancreatic (49.5%) cancer registrations in 2009 had a known anatomical subsite, while this figure was 87.3% for biliary cancer.
- Over 80% of oesophageal, stomach, duodenal and biliary cancer registrations had known morphology information in 2009. However, only 45.6% of pancreatic cancer had a known morphology. A large proportion of this unknown group was carcinoma, not otherwise specified (ICD-O-2 code 8010).
- In 2009, over 85% of all cancer types had a linked HES record. The proportion of HES linked records increased over time.
- Over 90% of oesophageal, stomach, duodenal and biliary cancer registrations in 2009 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 positive) was poor, although in some cases there was an increase in the proportion with a valid known record over time.

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 2000-2009. 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.

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/>

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.