



Public Health
England

Data Quality

Do we believe our data?

NCIN Central Nervous System Workshop
26th March 2014

Sally Vernon

Head of Quality and Analysis, National Cancer Registration Service

sally.vernon@ecric.nhs.uk

Overview

- What's good about our data?
- What are the problems with our data?
- What are we doing to fix them?

Strengths of the data

- Population level cancer data covering the whole country
 - Some countries only register a sample
 - Population based registration since 1960s
 - Population registration reduces bias / positive sampling of cancer cases

Strengths of the data

- Data feeds from all NHS Trusts
 - Standard dataset
 - Data liaison officers
 - Feedback to Trusts about data quality
- Wide variety of data sources used to compile case
 - Pathological data
 - PAS data
 - And more!
 - MDT data
 - Death data

Strengths of the data

- Data processed by specialist staff
 - Registration officers focus on cancer only
 - Detailed training programme to understand cancer sites and coding systems
 - Specialist QA officers develop data validations, checks and reports
 - Clinical engagement

Strengths of the data

- Good links with other datasets
 - Hospital Episode Statistics
 - Radiotherapy Dataset
 - Systematic Anti Cancer Therapy Dataset
 - Mortality data
 - PROMs and Patient Experience Survey
 - Audit data
 - CWT etc...

Strengths of the data

- Wide dataset collected
 - Over 490 data items specified in COSD
 - Site specific data items developed working with clinicians
 - Data covers broad variety of topics from biomarkers to treatment to patient experience!

Weaknesses - Timescales

2010	COSD not mandated SACT did not exist Registries using regional systems
2011	COSD not mandated First registries migrated to national system Still waiting for full linkages to HES, RTDS etc
2012	COSD not mandated Majority of registries on national system but not all Still waiting for full linkages to HES, RTDS etc
2013	COSD core data items mandated All registrations on national system Year not yet fully processed December diagnosis cases only had first three months of treatment
2014	COSD core and site specific data items mandated All registrations on national system Only three months into the year!

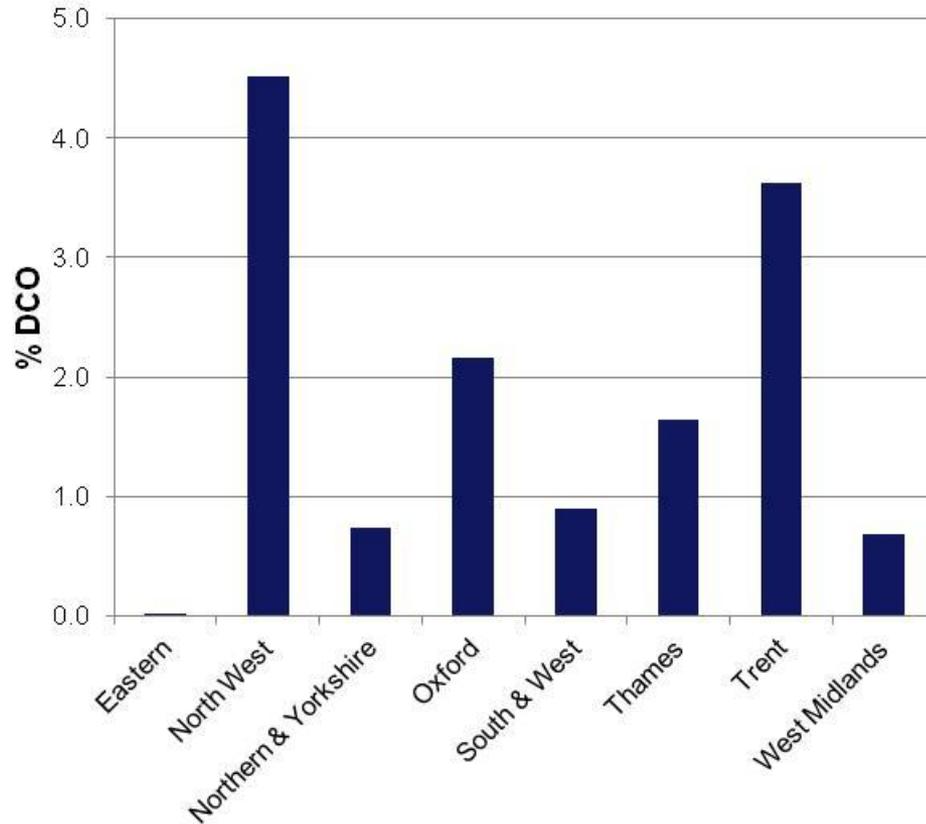
Need older data

- May be the most recent linked dataset
- Rare cancers or small geographies may need multiple years of data to have significant numbers of cases for analysis
- Can't use five year survival as outcome measure if five years has not passed!
- But older data do not have the full dataset that will be available for 2013

Weaknesses – UK data

- Work with Wales, Scotland, Northern Ireland and Ireland in partnership as the UKIACR
- But although attempts are made to standardise, variation can occur
- Routine datasets available for English data (eg HES, CWT, audit data) often not available for celtic countries

Weaknesses – Death Certificate Only



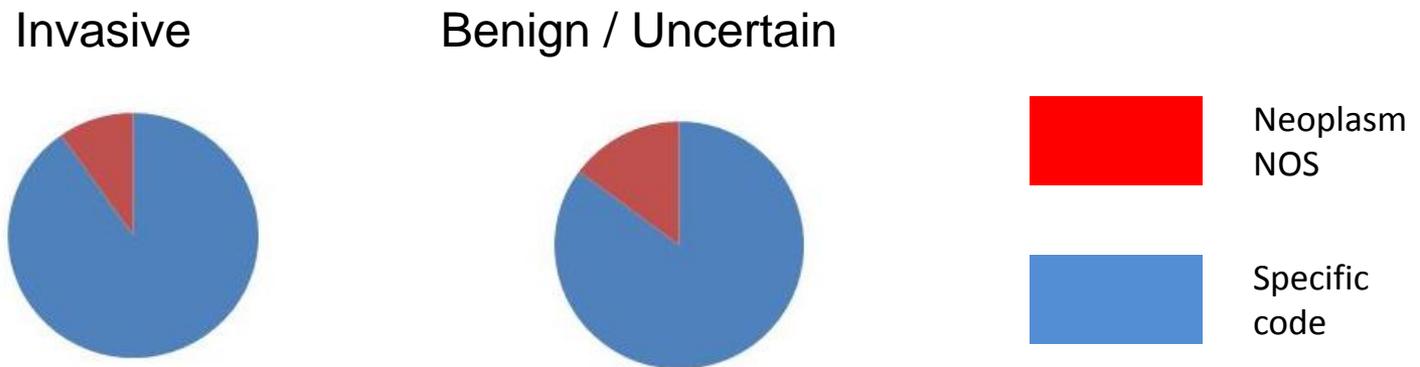
- Cases where only evidence is a death certificate
- Historically wide variation between offices
- Excluded from survival
- May indicate missing of live cases

Weaknesses – Clinical diagnosis

- Historically, cancer registries focused on pathological data
- Clinically diagnosed cases may be missed
 - Or reported as Death Certificates and excluded from survival
- Improving now
 - Access to MDT data
 - Imaging data

Weaknesses – Bucket codes

- Different brain cancers have very different care pathways and outcomes
- Cannot identify type of brain cancer without good morphological coding
- Historically, many brain cancers have been given bucket diagnoses



2005 tumours – over 1 in 10 coded as Neoplasm NOS

Weaknesses – Non-invasive tumours

- All brain tumours are a registrable condition
- But national statistics have historically focused on invasive tumours
- When the data are not being used, hard to identify data quality issues
 - One regional registry stopped submitting D32 (benign neoplasm of meninges) to ONS for over 10 years, and this wasn't spotted as no-one was analysing the data!
- Reported incidence rates of pituitary tumours strongly depend on
 - a) Amount of imaging being done, leading to incidental findings
 - b) Access of cancer registries to imaging data – better data, higher incidence rate

Weaknesses – Private hospitals

- The National Cancer Registration Service works with all hospitals treating cancer patients, and encourages them to submit data
- However, private hospitals treating private patients do not have same mandation as NHS hospitals
 - NCRS encourage them to submit data, and feed back reports to encourage them
 - But engagement with these hospitals with very small caseloads not always a priority
 - Will not get full COSD / HES / CWT data from these hospitals.

Weaknesses – Diagnosis date

- International definitions of the formal date of diagnosis
- Very focused on pathological confirmation
 - If no pathological confirmation, can use date of imaging etc
- Consequentially, some biases
 - The act of taking a biopsy can reduce measured survival time
- Particularly a problem for very short term survival cancers often diagnosed using imaging – eg glioblastomas!

Weaknesses – Lymphomas

- Historically, cancer registries coded in ICD 10
- This coded that a lymphoma had been diagnosed, but not where in the body
- No way to identify lymphomas of the brain in historical data
- ICD O3 solves this problem

Weaknesses – Metastases

- Primary brain cancers only part of workload
 - Metastases of other primary cancers to the brain are a significant proportion of all tumours in the brain
- National data on metastases historically poor
 - Good at saying 'it has metastasized' but not 'where to'
- Reviewing the data we collect on recurrence and metastases now COSD data is being collected.

Conclusions

- National Cancer Registration Data is a great resource – we have a world leading data set that allows us to understand and improve patient care across the country.
- There are known weaknesses in the available cancer data. It is important to consider these when planning work and interpreting analyses
- Recent developments such as one English National Cancer Registration Service, COSD and SACT will hugely improve data quality
 - but if we are looking at outcomes like 5 year survival there is a delay before these data are available.
- Improving and using data about cancer patients helps to drive up standards of cancer care. But we must understand the data quality issues of the data we use, to understand the findings we get.