

Rapid cancer registration in England created in response to the COVID pandemic

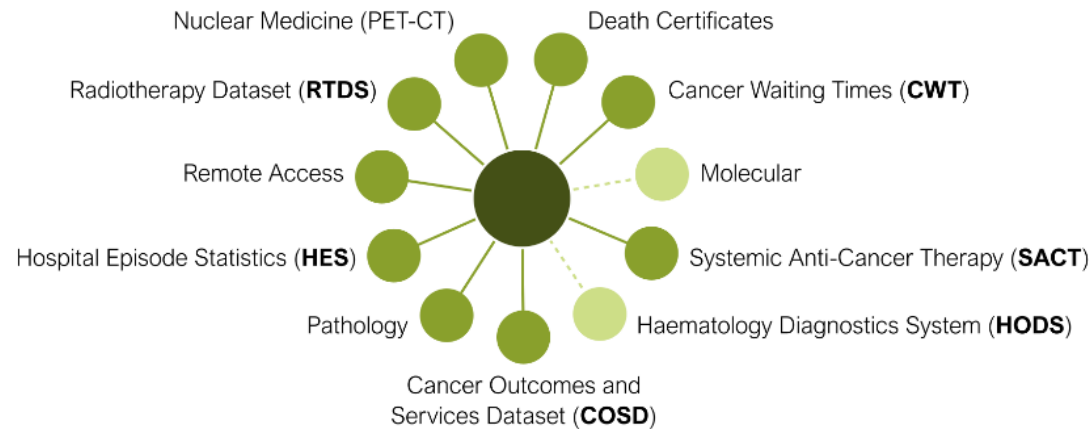
Sean McPhail, Carolynn Gildea, Kate Fleming
National Disease Registration Service, NHS England

Overview: What is the English Rapid Cancer Registration Dataset

- Abstract - the normal and rapid cancer registration datasets
- Rationale for creating the RCRD
- Build process for the RCRD
- Data quality – incidence and stage
- Use and outputs

Abstract – normal and rapid English cancer registration data

National Cancer Registration Data (**NCRD**),
'Gold Standard' data

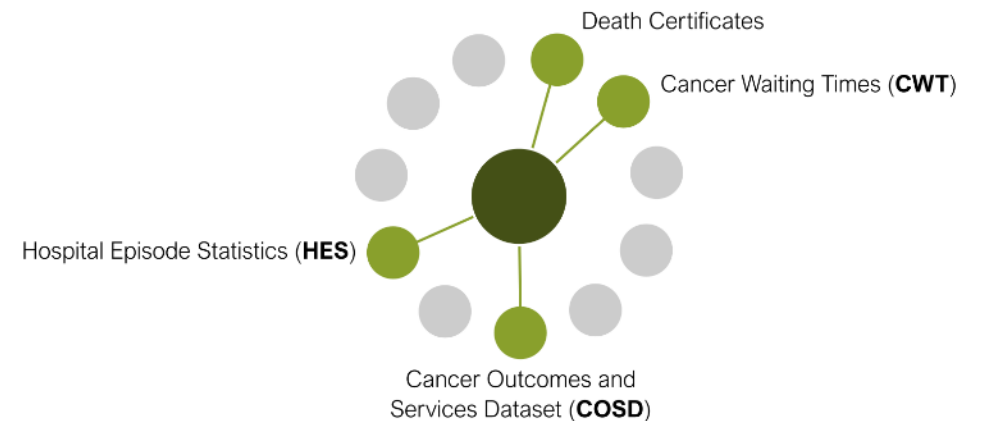


Available 21 months post-diagnosis: full reporting, all clinical detail

Available from January 1995 (in ICD-10)

Staging for all stageable cancers

Rapid Registration (**RCRD**)



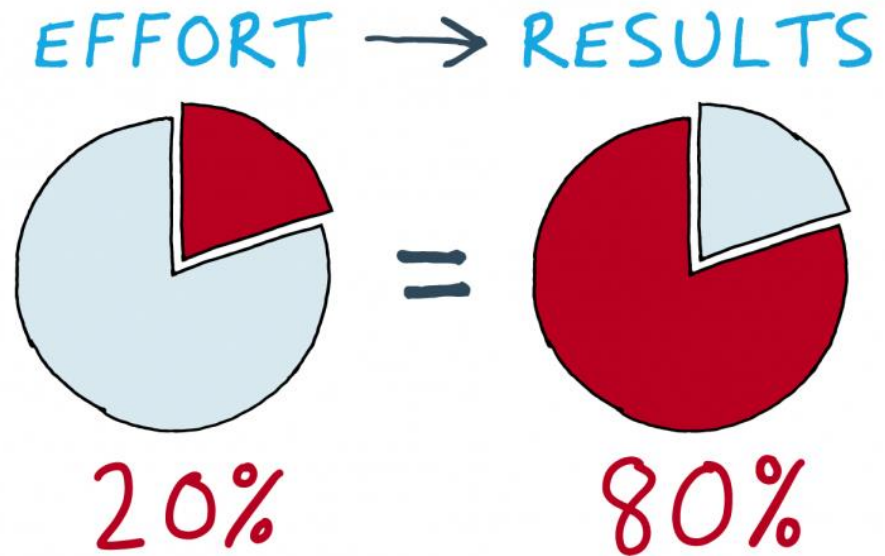
Available 4 months post-diagnosis: rapid monitoring of case numbers

Available from January 2018

Staging for 13 cancer types

Rapid Cancer Registrations: Rationale

THE PARETO PRINCIPLE



LEAD
50 models for success

Rapid Registration tumours: inclusion criteria

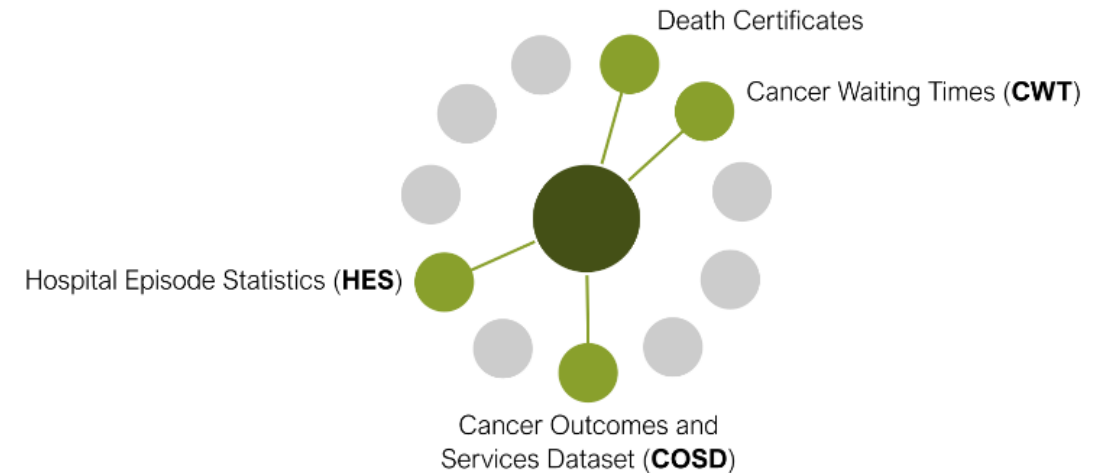
1. Created for 2018-01-01 to most recently available data
2. Covers all malignant cancers and selected non-malignant cancers: Breast, Bladder & Brain (ICD-10 C00-C97, plus D05, D09, D32, D33, D35, D41, D42-D44)
3. Persons appearing in English data – includes cross-border cases.
4. Staged (malignant) sites: Bladder; Breast; Colorectum; Kidney; Lung; Lymphoma; Melanoma; Oesophagus; Ovary; Pancreas; Prostate; Stomach; Uterus

Monthly build process for the RCRD dataset

Arriving at NDRS within 4 months:

- Cancer Outcomes & Services Dataset – 91.1%
(Hospital cancer patient management systems)
- Hospital Episode Statistics – 3.7%
(Hospital patient administration systems)
- Cancer Waiting Times – 1.5%
(Hospital cancer referral speed tracking systems)
- Mortality data – 3.7%
(National death certificate information)

Rapid Registration (**RCRD**)



Monthly build process for the RCRD dataset

1. Drop/archive the earlier version of the dataset.
2. Summarise everything useful from the four incoming datasets into an unified events 'pathway' table.
3. Refine, combine and filter those events into rapid "proxy" tumour diagnoses.
4. (Number of patients in pathway table >> patients in tumour table.)
5. Add auxiliary information to bare tumour & patient detail (in total ~9,600 lines of SQL code, ~3700 lines of .rmd data quality assessment).

Rapid Registration structure

RAPID_PATHWAY

AVPID
INDIVIDUALID
PATIENTID
NHSNUMBER
SOURCE_TABLE
SOURCE ID

EVENT_TYPE
EVENT_PROPERTY_1
EVENT_PROPERTY_2
EVENT_PROPERTY_3
EVENT_DATE
EVENT_END
TRUST_CODE

**34 event types
summarising 4 data types**

ID fields

Tumour fields

“What, where & when”

Patient fields

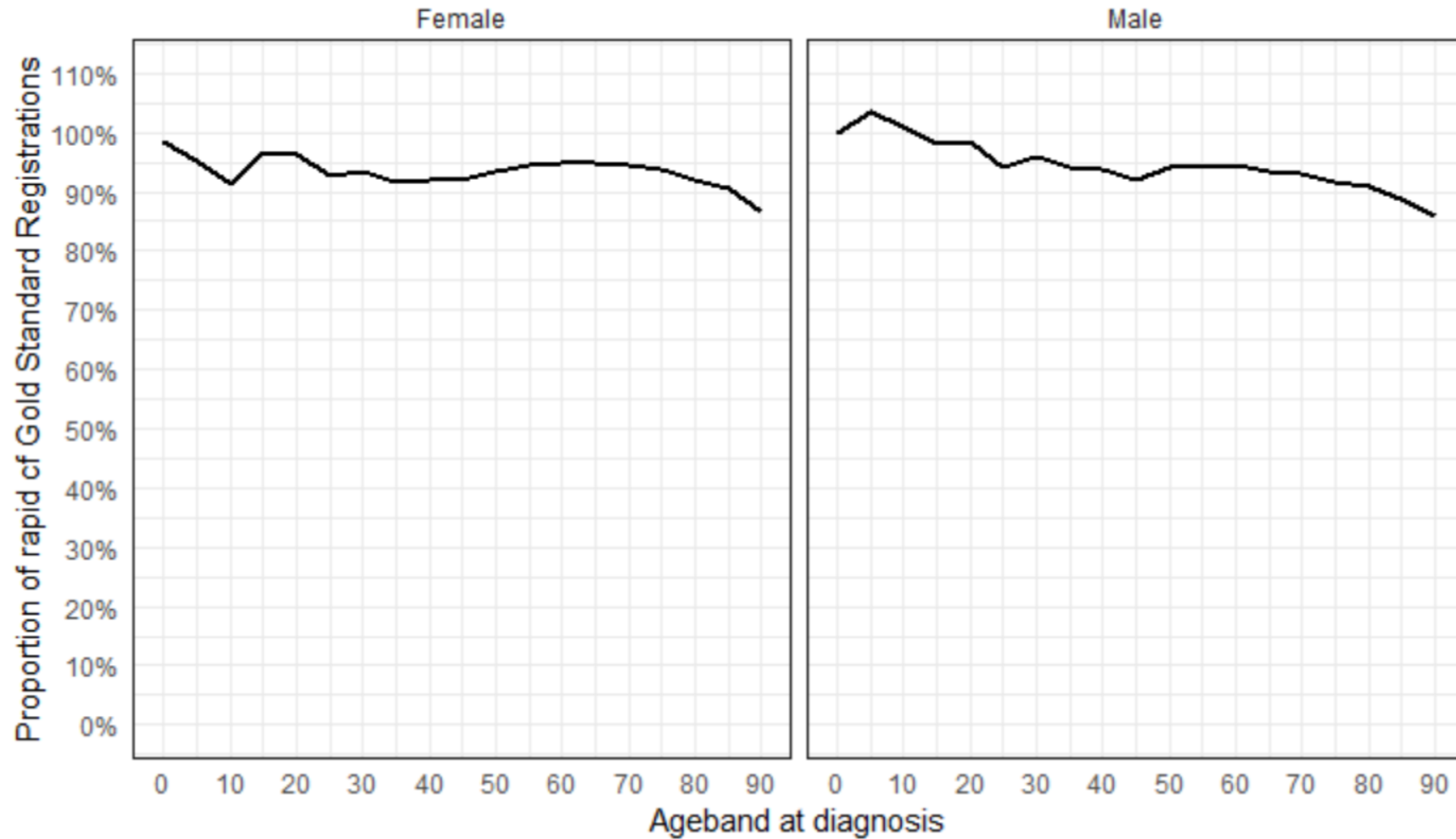
RAPID_TUMOUR

TUMOUR_AVPID
INDIVIDUALID
PATIENTID
NHSNUMBER

DIAGNOSIS DATE
BASIS OF DIAGNOSIS
TUMOUR SITE
TUMOUR MORPHOLOGY
ROUTE TO DIAGNOSIS (HES)
CHARLSON COMORBIDITY
STAGE

BIRTH DATE
SEX
POSTCODE
SURNAME
FORENAME
ETHNICITY

Comparing to normal cancer registration data, 2018 cases (nmisc excluded)

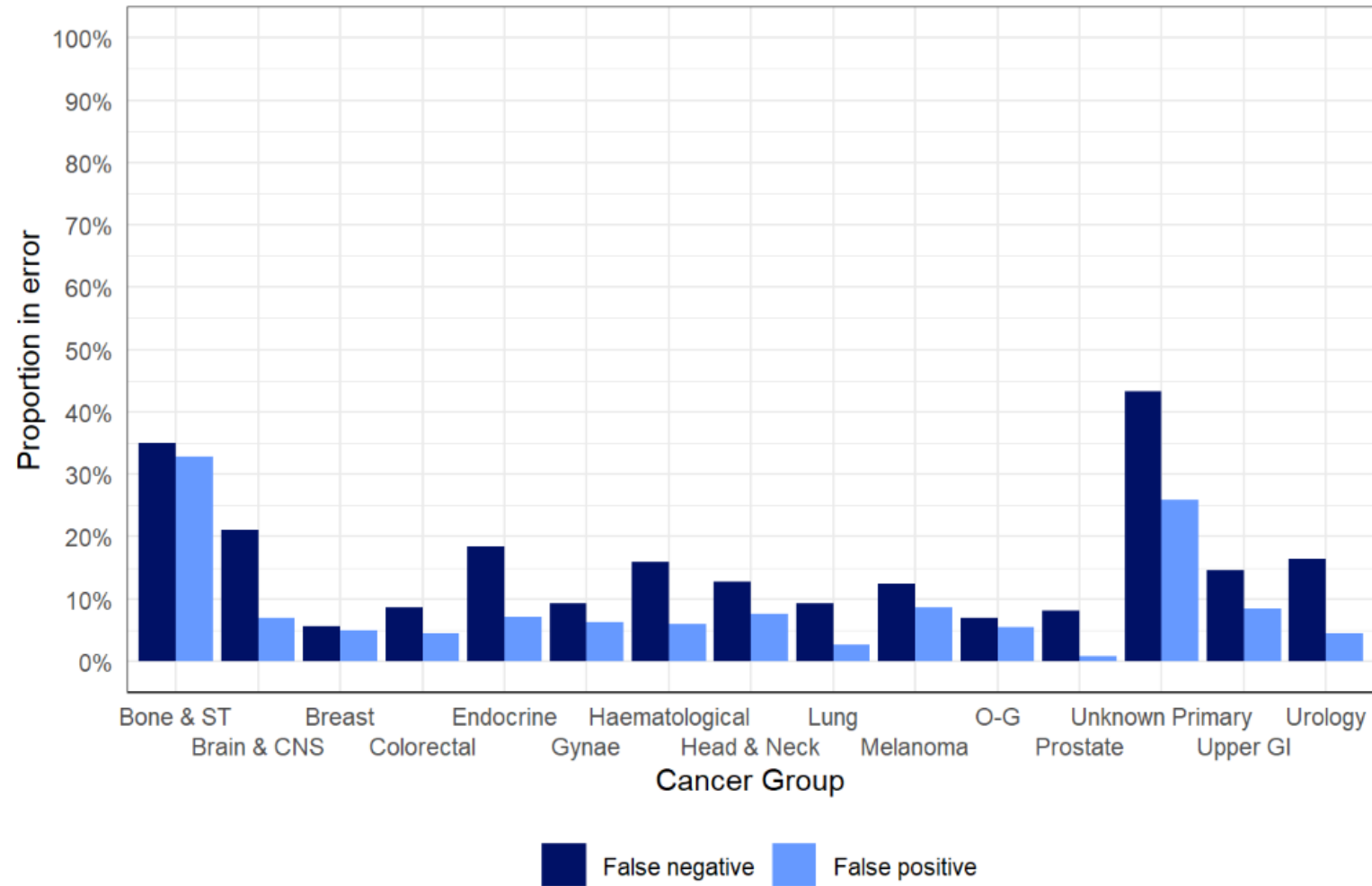


Source: NHS England, National Cancer Registration and Analysis Service

Further defining quality – False Positives, False Negatives

- Compare rapid registration data to normal registrations, Apr 2018 to Sept 2018.
- Define a matched/successful proxy registration as:
 1. Matching in a broad cancer group (e.g., C18-C21, or C50 & D05)
 2. Matching diagnosis date to within 3 months
- Errors:
 - **False negative (FNE)** – real registration but no proxy registration (“*missing data*”)
 - **False positive (FPE)** – no real registration but proxy registration (“*bad data*”)

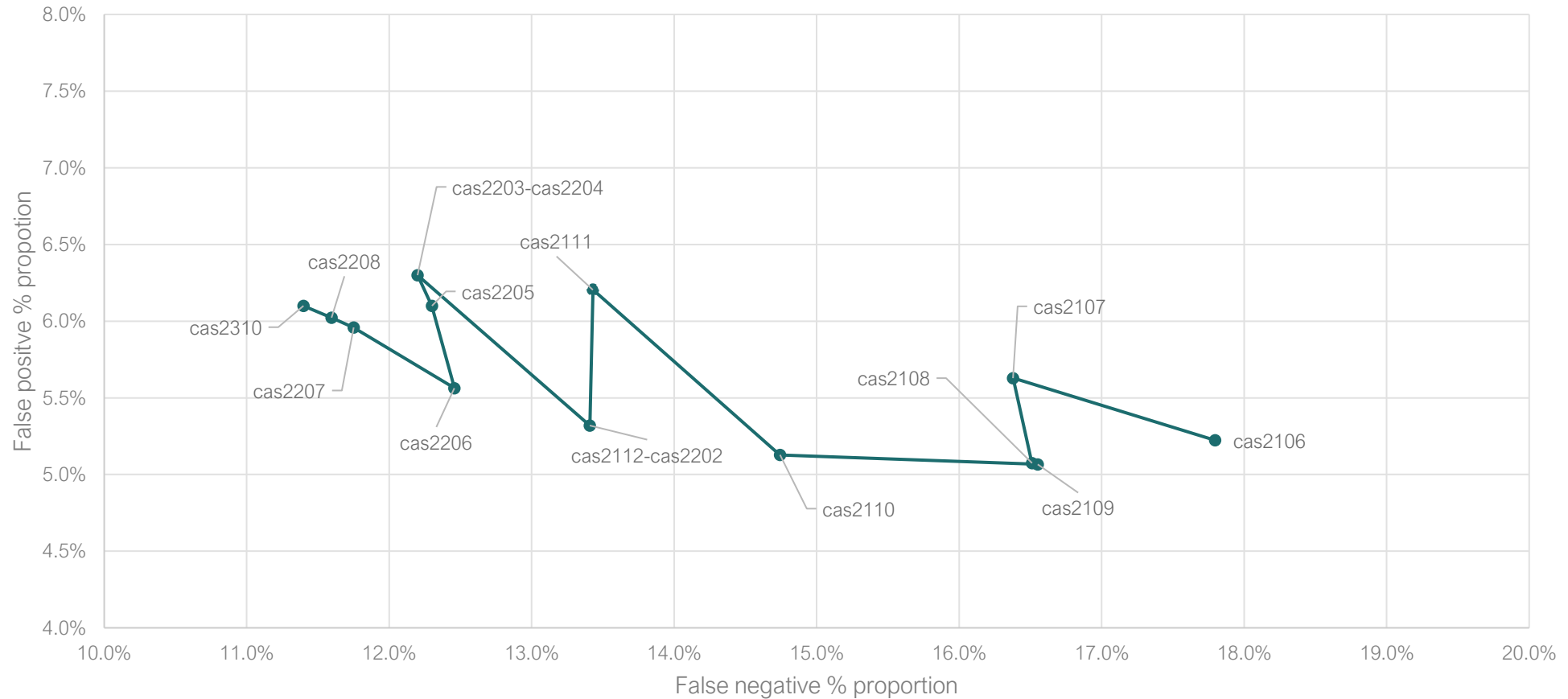
False Positives (6.1%), False Negatives (11.4%)



CNS: Central Nervous System; GI: Gastrointestinal; O-G: Oesophagogastric; ST: Soft Tissue
Source: NHS England, National Cancer Registration and Analysis Service

Trend in FNE and FPE (June 2021 – August 2022)

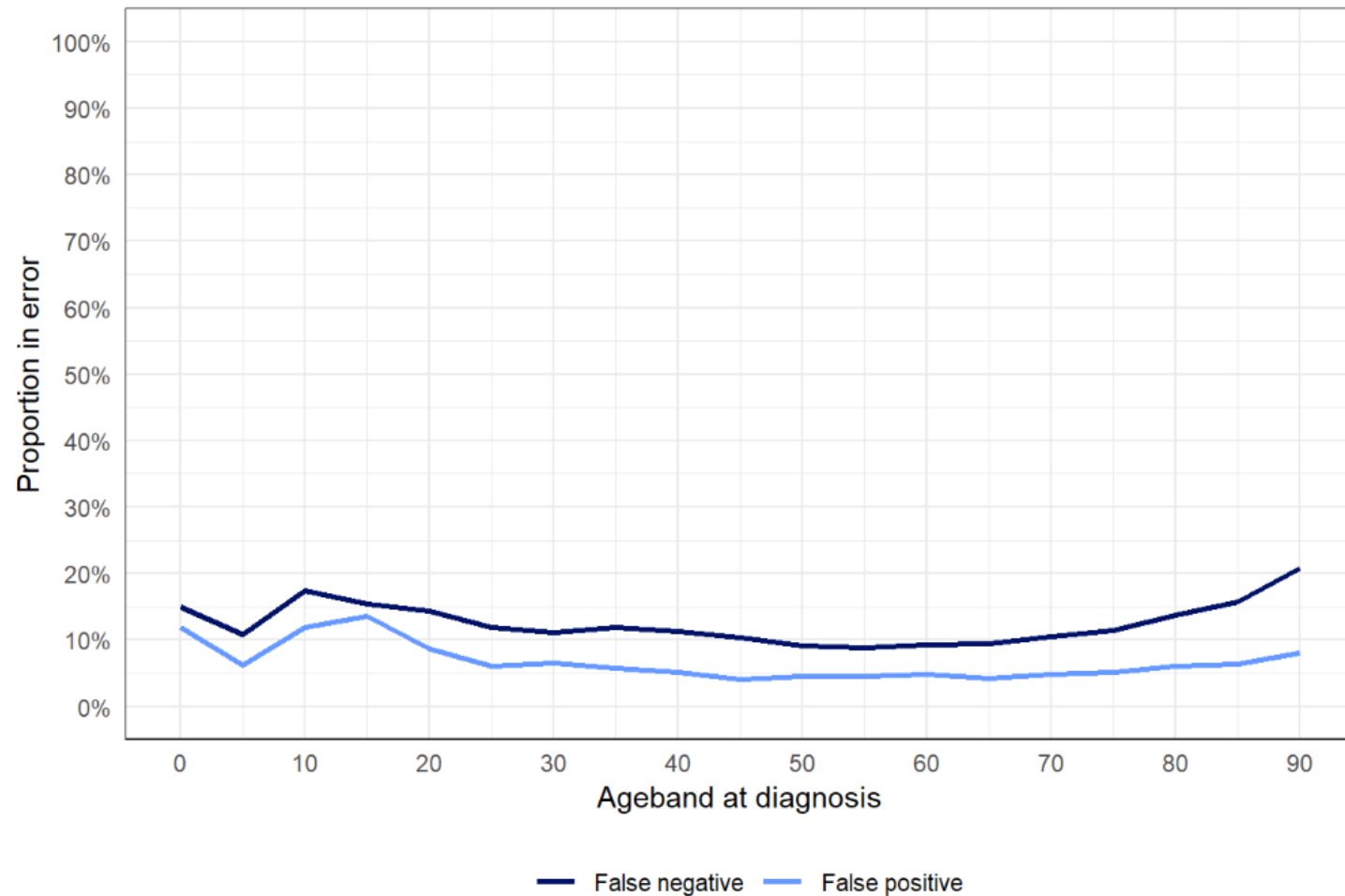
RCRD False Negative and False Positive error proportions by snapshot



Monthly snapshots coded as: casYYMM

What are the False negatives?

- Cancers not represented in COSD...
- Age dependence is explained by more clinical pathways in older people, more early fatality, lighter 'contact' with the system.



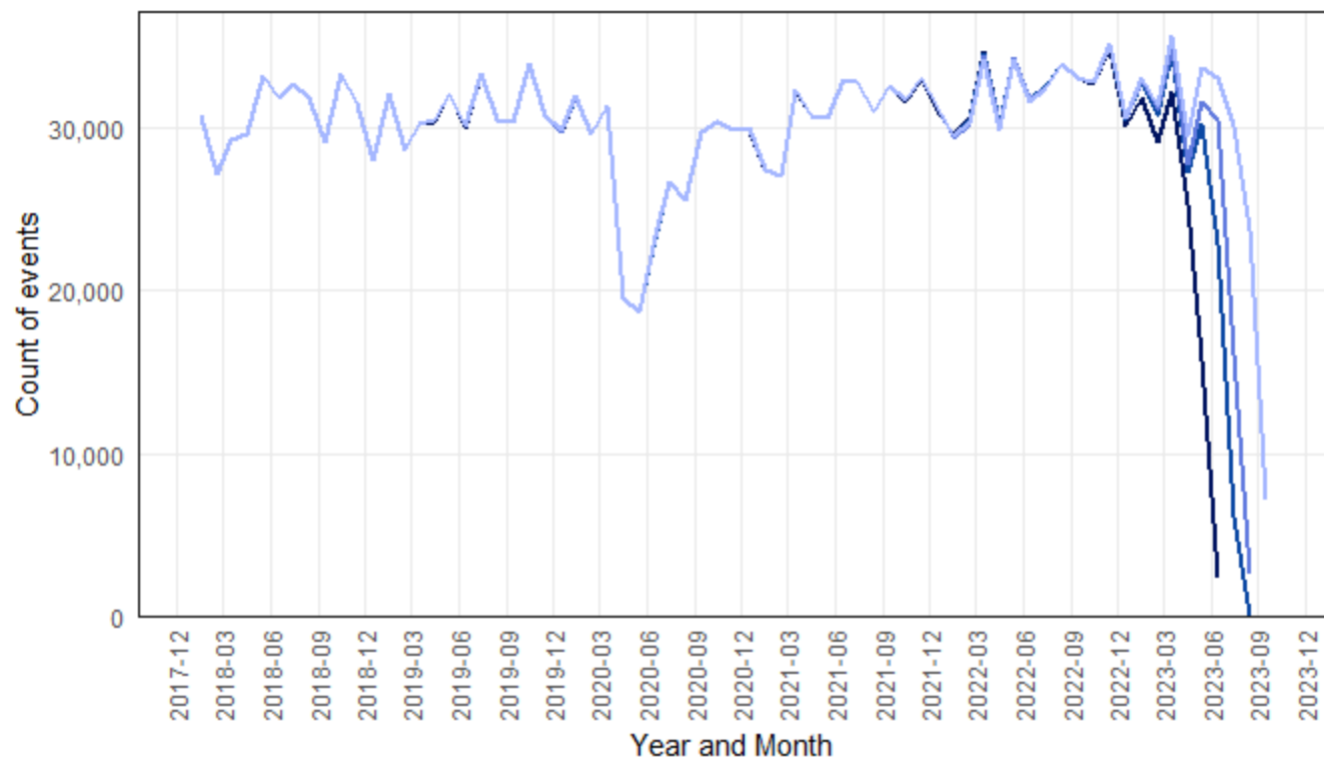
Source: NHS England, National Cancer Registration and Analysis Service

What are the False Positives?

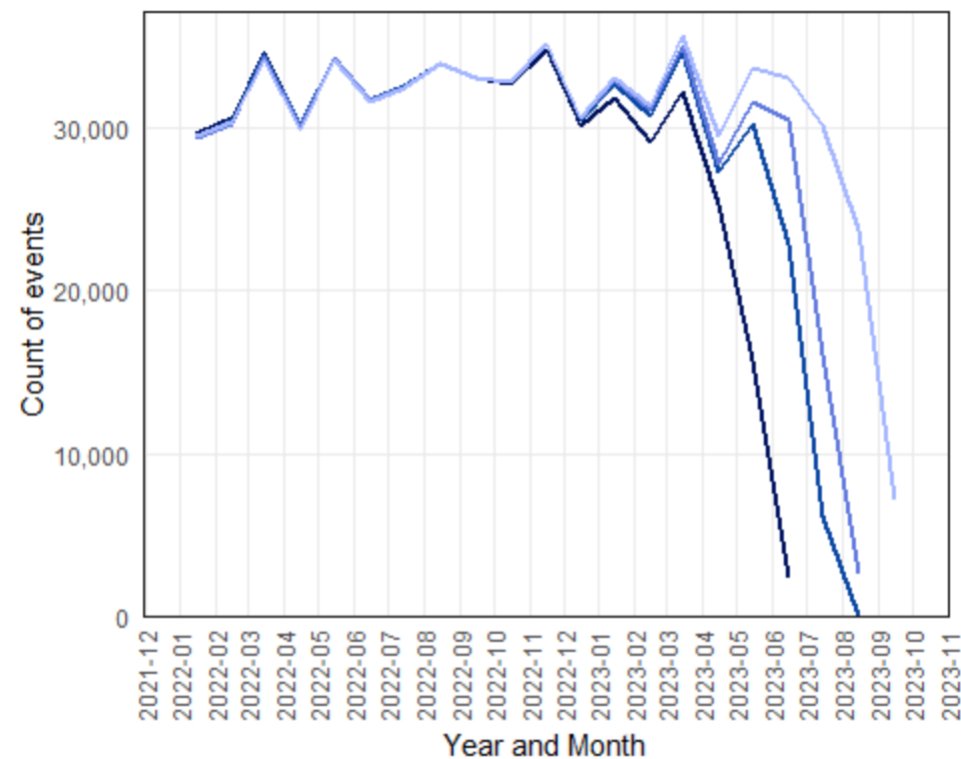
Not so clear cut, but approximately:

| | |
|-------|--|
| 30% | 'Tricky' tumour sites – (Unknown Primary/Sarcoma) |
| 20% | D-code / C-code mismatch in COSD and normal registration |
| 12.5% | COSD tumours proved mistaken and deleted |
| 5% | Melanoma / NMSC confusion |
| 32.5% | Tumours diagnosed in previous years reoccurring |

'snapshot creep': accumulating incidence counts



— cas2307 — cas2308 — cas2309 — cas2310

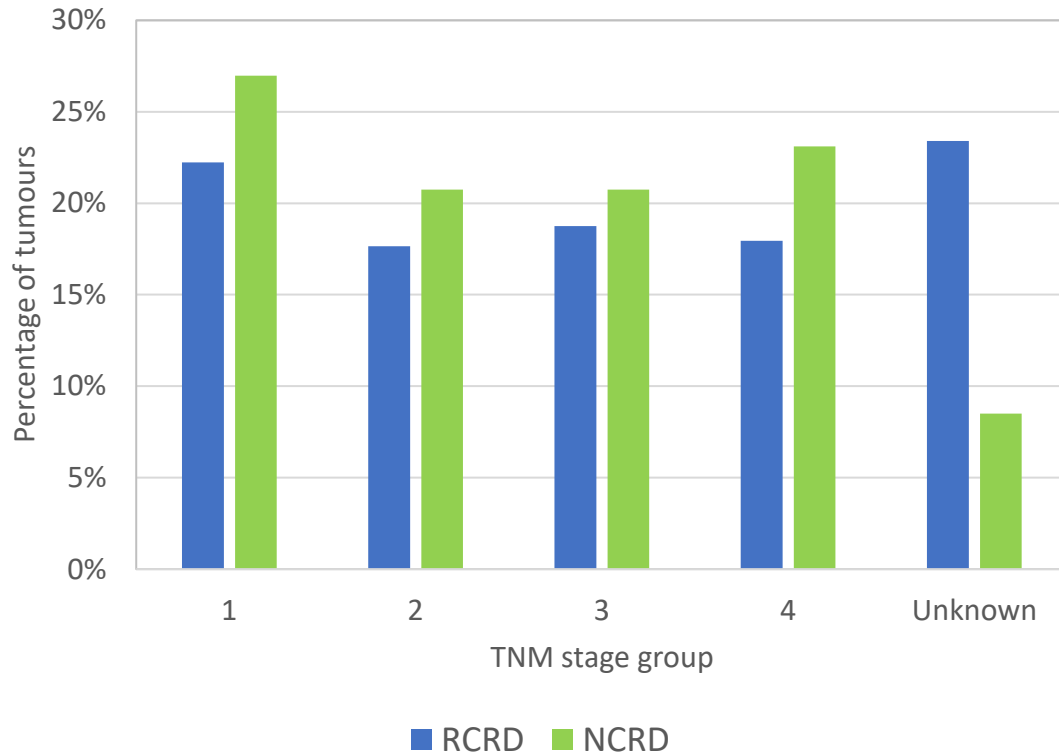


— cas2307 — cas2308 — cas2309 — cas2310

- Monthly quality review needed – publication threshold of 90% completeness
- Typically publish within 4 months of diagnosis: data on patients diagnosed January published at the end of April.

Stage at diagnosis (Big 4 cancers) – TNM stage

Stage proportion by TNM Group - Breast, Colorectal, Lung and Prostate



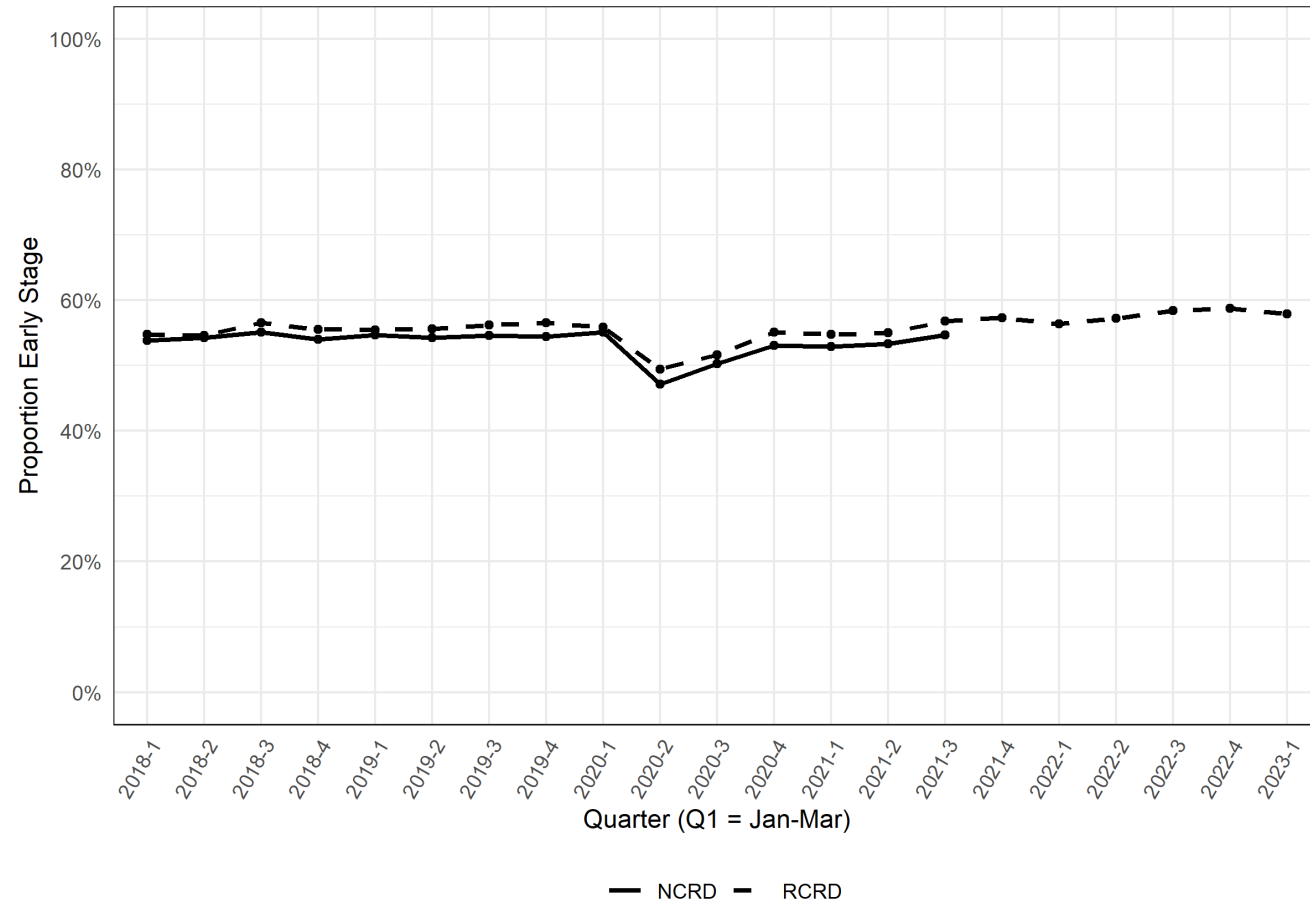
“~77% of (Big 4) cancers are staged, of those 80-90% are correct.”

Proxy Stage

| Normal stage | 1 | 2 | 3 | 4 | Unk |
|--------------|------|------|------|------|------|
| 1 | 89% | 9% | 2% | 1% | 30% |
| 2 | 6% | 83% | 5% | 1% | 16% |
| 3 | 3% | 5% | 85% | 3% | 12% |
| 4 | 1% | 2% | 6% | 95% | 16% |
| Unknown | 2% | 2% | 2% | 1% | 26% |
| | 100% | 100% | 100% | 100% | 100% |

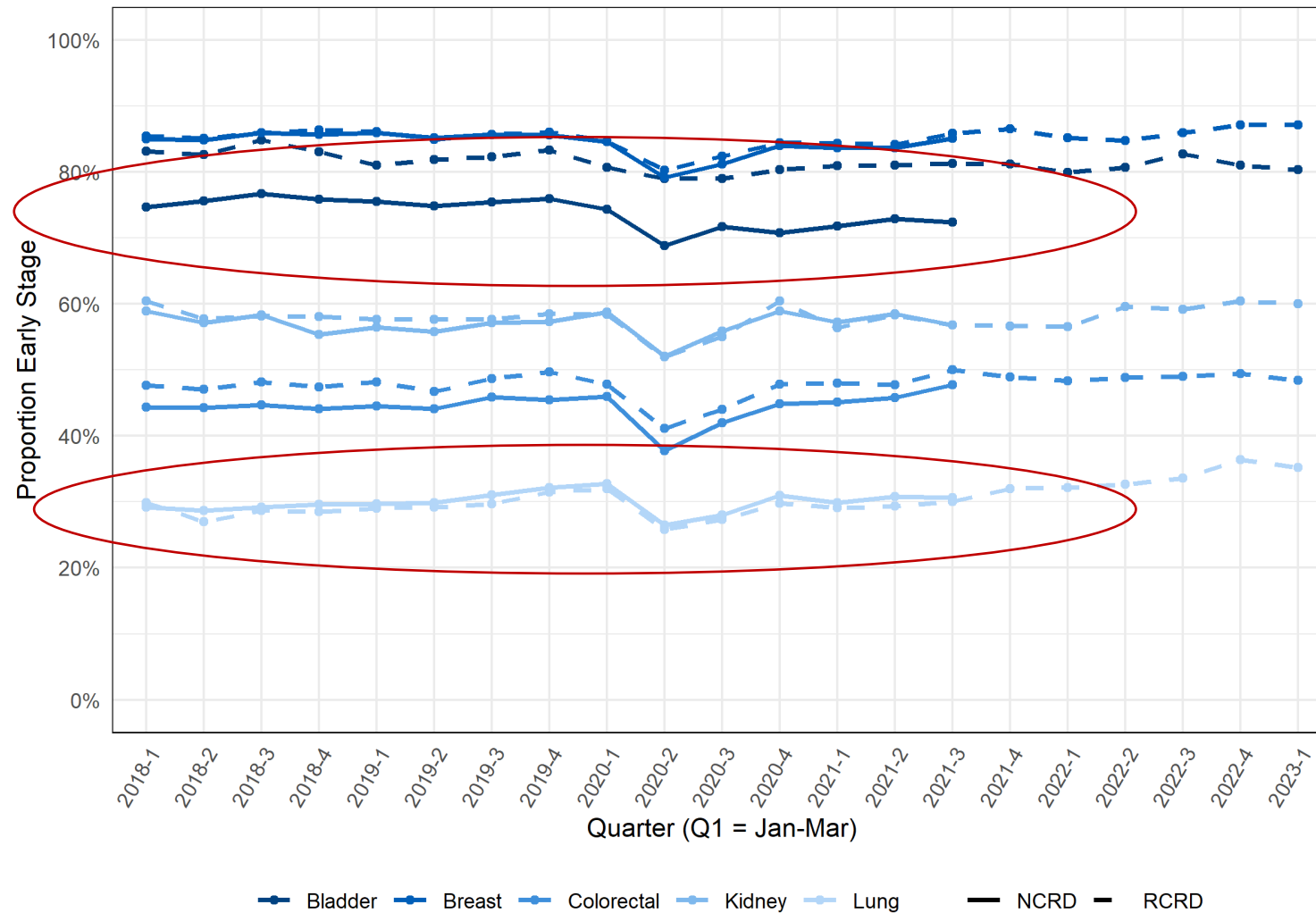
| Normal stage | Early | Late | Unk |
|--------------|-------|------|------|
| Early | 93% | 5% | 46% |
| Late | 5% | 94% | 28% |
| Unknown | 2% | 2% | 26% |
| | 100% | 100% | 100% |

Overall comparison of proportion early stage in RCRD (dashes) to NCRD (solid)



Source: NHS England, National Disease Registration Service
2021-1 to 2021-3 is provisional data for Jan-Mar 2021 in NCRD

Proportion of early stage: variation by cancer type



Source: NHS England, National Disease Registration Service
2021-1 to 2021-3 is provisional data for Jan-Mar 2021 in NCRD

Data online

COVID-19 rapid cancer registration and treatment data

Introduction Uses and Limitations Update log Summary **Demographic factors** Time to treatment - Demographics Geography

Select Cancer group: All sites combined

Select demographic level: Route to Diagnosis

Time trend (chart) Time trend (table) Time trend (factor proportions) Annual comparison (chart) Annual comparison (table)

Annual comparison (factor proportions) Downloads

New cancer diagnoses, England, January 2019 to March 2023
Cancer group: All sites combined

Route to Diagnosis

- Emergency presentation
- GP referral
- Other
- Other outpatient
- Screening
- TWW

This work has been produced by the National Disease Registration Service (CAS2309).

COVID-19 rapid cancer registration and treatment data

Introduction Uses and Limitations Update log Summary Demographic factors **Time to treatment - Demographics** Geography

Tumour resection Chemotherapy Radiotherapy Proportion of treatment (Tumour resection) Proportion of treatment (Chemotherapy)

Proportion of treatment (Radiotherapy) Downloads

Select Cancer group: All sites combined

Select demographic level: Total

Using the rapid cancer registration data, the tool presents treatment proportion data by time to treatment, for England, from January 2018 onwards:

- Tumour resection proportion: Proportion of cancer diagnoses receiving a tumour resection, by diagnosis date
- Chemotherapy proportion: Proportion of cancer diagnoses receiving chemotherapy, by diagnosis date

Tumour resection treatments by time to treatment, England, January 2019 to March 2023
Cancer group: All sites combined

Total

Percentage of diagnoses receiving a tumour resection

This work has been produced by the National Disease Registration Service (CAS2309).

Publicly accessible data: <https://www.cancerdata.nhs.uk/covid-19/rcrd>

Documentation: <https://digital.nhs.uk/ndrs/data/data-sets/rcrd>

Apply for low level data: <https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue/ndrs-rapid-cancer-registrations-data-set-rcrd>

Summary: capabilities of the Rapid Reg. dataset

The Good

- Reasonable proxy for cancer registrations (where data overlaps).
- Typically 4 month delay (varying with data type, and allowing for 'creep').
- Strong potential to link to other datasets.

The less good

- Some systematic biases in FNE (missing registrations) in particular, e.g., age.
- Variation in staging agreement for some cancer types.

Many thanks to:

Tom Bacon

Jackie Charman

Cong Chen

Sophie Jose

Ravneet Sandhu

Sally Vernon

Kirstin Roberts

Hanhua Liu

Peter Jones

Any Questions?

Sean McPhail

Principal Analyst

sean.mcphail1@nhs.net

This work uses data that has been provided by patients and collected by the NHS as part of their care and support. The data are collated, maintained and quality assured by the National Cancer Registration and Analysis Service, which is part of NHS England.