Public Health England

Protecting and improving the nation's health

Cancer Outcome and Services Data set

Specification

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-leading science, research, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Published September 2019
PHE publications gateway number: GW-621

PHE supports the UN Sustainable Development Goals

Data Coordination Board

This information standard (DCB1521) has been approved for publication by the Department of Health and Social Care under section 250 of the Health and Social Care Act 2012.

Assurance that this information standard meets the requirements of the Act and is appropriate for the use specified in the specification document has been provided by the Data Coordination Board (DCB), a sub-group of the Digital Delivery Board.

This information standard comprises the following documents:

- Specification
- Implementation guide
- Change request

An Information Standards Notice (DCB1521 Amd 13/2019) has been issued as a notification of use and implementation timescales. Please read this alongside the documents for the standard.

The controlled versions of these documents can be found on the NHS Digital website. Any copies held outside of that area, in whatever format (e.g. paper, email attachment), are considered to have passed out of control and should be checked for currency and validity.

Date of publication: 12 September 2019.

Contents

Title	Cancer Outcome and Services Data set (COSD) - Specification		
DCB Reference	DCB1521 Amd 13/2019		
Sponsor	Dr Jem Rashbass	**Status**	Final
Senior Responsible Officer	Professor John Newton	**Versions**	COSD v9.0 Pathology v4.0
Developer	Andrew Murphy		
Author(s)	Public Health England	**Version Date**	06/09/2019

Amendment history

Version(s)	Date	Amendment History
COSD v9.0 & Pathology v4.0	20 June 2019	Draft document sent to DCB for initial review
COSD v9.0 & Pathology v4.0	4 July 2019	Final amended document sent to DCB for board review
COSD v9.0 & Pathology v4.0	1 September 2019	Final amended document for publication

Approvals

This document has been approved by the following:

Name	Signature	Title / Responsibility	Date	Version
COSD Advisory Board	COSD Advisory Board	Cross Organisation Board	4 April 2019	v9.0 & v4.0
COSD Governance Board	COSD Governance Board	Cross Organisation Board, reporting to Professor Chris Harrison, National Clinical Director (NHS England) and Dr Jem Rashbass, Director for National Disease Registration (Public Health England)	30 April 2019	v9.0 & v4.0

Executive summary

The purpose of this document is to provide guidance intended to support providers of Cancer Services and developers (both in-house and commercial system suppliers), to prepare for the implementation of the Cancer Outcomes and Services Data set COSD v9.0 and v4.0 for Pathology from April 2020.

All documents (or links to them) can be found on either the COSD[1] or NHS Digital[2] websites unless otherwise stated. These provide assurances that the proposed approach meets the business requirements identified in the requirements specification for DCB1521 Amd 13/2019 have been adequately researched and can be delivered.

This is an update to an existing information standard DCB1521 Amd 74/2016 and is required to ensure that the data still meets the business objectives, scope and content of the standard and continues to be clinically accurate and relevant.

In order to maintain the clinical accuracy, it is important to regularly review COSD with clinical experts from across the NHS, including analysts and NHS England.

Occasionally other information standards have specific data items which interact with COSD. Where this happens, additional consultation with the developers of those standards was undertaken to ensure all data items remain accurate and are updated where necessary.

Although in the most part all changes are described as COSD, there is a separate pathology data set, which requires a different schema pack due to the different linkage required. This will be referred to within this document as COSD Pathology data set v4.0.

It is important to note that pathology is at v4.0 due to this being the fourth schema required for this data set, and that pathology was first mandated as a separate data set in 2016. All pathology data items have now been removed from the main COSD v9.0 data set and can only be submitted via the pathology data set and its set of schemas.

---

[1] http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd
[2] http://www.digital.nhs.uk/isce/publication/dcb1521

Introduction

Background

The Cancer Outcomes and Services Data Set (COSD) is the national standard for reporting cancer for the NHS in England. The National Cancer Registration and Analysis Service (NCRAS) are responsible for ongoing maintenance, development and implementation. The data sets relate to all cancer patients, both adult and paediatric, in acute inpatient and outpatient settings, but does not include private patients or primary care.

The Cancer Outcomes and Services Data set (COSD) are compiled data sets, which provides the standard for secondary uses. The standard consists of:

- a set of individual data items, with their definitions
- the assemblage of these data items into discrete data sets
- the means of flowing the data items
- compilation of the data items into 2 reconciled and verified data sets

Overtime clinical data items may have changed or been amended by internationally recognised bodies e.g. Royal College of Pathologists or international staging systems (revised international prognostic scoring system for myelodysplasia). These must be acknowledged, and amendments made.

The Standard supports national statistics, allowing local and national comparisons of performance and service activity. Additionally, the output supports commissioning and service development through the provision of relevant information on service delivery and outcomes.

COSD has removed the interdependency between the National Cancer Waiting Times Monitoring Data set (NCWTMDS). In v9.0 developers from NHS England and Public Health England (PHE), have worked together to align both data sets, to reduce wherever possible the burden of data collection.

Data are collated via the NCRAS local offices, and formal mechanisms for transmission of data from Providers to registries have been extended to carry the COSD data sets.

Who does COSD apply to?

COSD applies to the following key groups and organisations:

- cancer centres, cancer units and all other providers of NHS commissioned cancer services
- developers and suppliers of electronic systems for use in NHS commissioned cancer centres and NHS provider services
- organisations purchasing electronic systems for use in NHS commissioned cancer centres and NHS provider services
- users of secondary data about cancer at both national and local levels, including:

At a national level:

- Department of Health and Social Care (DHSC)
- National Cancer Registration and Analysis Service
- Public Health England
- NHS England and NHS Improvement
- NHS Digital
- Care Quality Commission (CQC)
- NHS Improvement[3]

At a local level:

- cancer alliances
- local cancer service provider networks
- local NCRAS offices
- commissioners and providers

As COSD is for secondary uses, there is no intention for this to be used by Primary Care or Private Hospitals.

Note: It is important to note however, that if a patient is on a NHS pathway but the treatment is carried out in a private hospital (due to capacity issues or at the request of the NHS Trust), these data must be collected and reported (within COSD) by the NHS Trust, as if the treatment was carried out by them.

[3] https://improvement.nhs.uk/

Summary of changes

These 2 new versions of COSD (v9.0 and v4.0 for pathology), complete the changes started in 2017 and further updated in 2018. These additional changes were required in order to make the data sets clinically accurate, fully aligned with the Royal College of Pathologists (RC Path) Core data sets and also meet the recommendations within the Achieving World-Class Cancer Outcomes, A Strategy for England 2015 to 2020 (Cancer Taskforce Report)[4]. This required changes to the standard which include:

- an improved mechanism for recording all 'Non Primary Cancer Pathway' to improve collection and data quality
- some relocation/restructuring of items within the data set
- new sections within the 'Core' to record diagnostic procedures and acute oncology
- the mandation of key data items to improve data quality and reduce the burden of data processing within NCRAS, as outlined on page 10
- the addition of key new data within the Breast and Upper GI site specific sections
- the addition of a prefixed 'p' to all pathology data items along with the removal of all pathology data items from the COSD v9.0 data set, this reduces the potential in future versions of 2 or more data items having the same item number
- update of COSD pathology v4.0 data items to align with the RC Path 'Core' data sets
- a revision of the current schema specification, in order to continue to meet the business objectives of the standard

New data items have been added after an extensive (5-month) consultation was conducted with 40 key stakeholder groups and over 150 clinical experts:

- site specific Expert Advisory Groups (EAGs) within PHE
- experts from within the National Cancer Registration and Analysis Service (NCRAS)
- clinical support and advice from the chair of the Royal College of Pathologists Working Group on Cancer Services
- cancer charities (Cancer Research UK, AMMF – the Cholangiocarcinoma Charity, Breast Cancer Care, Living Beyond Cancer and MacMillan)
- patient groups and individuals through 'Use My Data' and the EAGs
- national cancer audits

[4] http://www.cancerresearchuk.org/about-us/cancer-strategy-in-england

- cancer and pathology system suppliers
- NHS England
- Cancer Waiting Times (CWT)
- quality surveillance
- cancer taskforce transformation board

This revision allowed the data sets to be clinically reviewed, validated and updated by experts in all fields of cancer, and provides a clinically sound set of data to be collected from April 2020 onwards.

However, it was recognised that there were still further improvements required within this release to improve data quality and to be more reflective of current and future changes in cancer treatments, outcomes and clinical care.

Mandating key data items was vital in this release to improve data quality and reduce the burden of data processing within NCRAS. If a treatment record for example is submitted without a date or modality, it is of no use as a cancer registration event, however a registration officer would still have to review each incomplete record submitted, this multiplied by thousands of incomplete records per year was an unsustainable practice.

The data set can now be effectively implemented to improve the collection of specific data items within each Trust, by using one of 2 data sets (depending on the department responsible for each data collection process):

- **COSD v9.0** – this is the data set which the cancer services teams need to collect excluding pathology, by removing the pathology data items from their workload this could reduce their burden of data collection by up to 30% across the whole data set
- **COSD Pathology v4.0** – from April 2020, pathology data items will only be able to be reported via the patholgy data set using the associated schema packs, this is mandated across all Trusts that supply these data in COSD XML directly from their pathology departments

Wherever possible duplication across the data set has now been removed and full explanations of how to collect these data within the new structure are provided within the change logs of each data set.

Finally, where there were data that are no longer part of a linked national data set (for example, the Royal College of Pathologists), these have also been removed from v4.0 of the COSD Pathology data set.

Implementation start and full conformance timeline

The following timeframe will be used to support the implementation, data collection and outline the full conformance dates:

- implementation will be between 6 September 2019 and 31 March 2020 (six-and-a-half months)
- data collection will start from 1 April 2020 (with a 3-month roll-out period between 1 April 2020 and 30 June 2020)
- full conformance from 1 July 2020 (reported in the July batch within the September upload)

Supporting documents

All the documents referred to in this specification document were submitted to the Data Standards Assurance Service (DSAS) for review under DCB1521 amendment Amd 13/2019.

Following acceptance by Data Coordination Board (DCB) and confirmation of authority to publish by the Department of Health and Social Care, the official Information Standards Notice (ISN) and related documents were published on 6 September 2019.

This Specification should be read in conjunction with the following documents, available at www.digital.nhs.uk/isce/publication/dcb1521:

- change request
- implementation guide
- information standards notice

http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd:

- COSD data set v9.0
- COSD v9.0 user guide
- COSD v9.0 technical guide
- COSD pathology data set v4.0
- COSD pathology v4.0 user guide
- COSD pathology v4.0 technical guide

https://isd.digital.nhs.uk/trud3/user/guest/group/0/home:

- COSD data set v9.0 schema pack
- COSD pathology data set v4.0 schema pack

These documents are intended to support providers and developers that wish to identify and plan changes to their systems. The standard will be formally issued via DCB as an approved standard and additional documents (for example the data sets, user guides and technical guides), will be available to download via the NCIN website, where a new page for downloads will be created.

Related standards

The following standards should also be read in conjunction with this information standard:

- DCB0147 – National Cancer Waiting Times Monitoring Data Set[5]
- SCCI0111 – Radiotherapy Data Set[6]
- DCB1533 – Systemic Anti-Cancer Therapy Data Set[7]
- SCCI1577 – Diagnostic Imaging Data set[8]
- SCCI0021 – International Classification of Diseases[9]
- SCCI0034 – SNOMED CT[10]
- DCB2094 – Sexual Orientation Monitoring[11]
- Royal College of Pathologists Standards and Data sets for Histopathology Reporting on Cancers and Tissue Pathways[12]

Contacts

The following are a list of key contacts responsible for the development and management of the data sets:

- sponsor – Dr Jem Rashbass (PHE)
- senior responsible officer – Professor John Newton (PHE)
- developer – Andrew Murphy (PHE)
- implementation manager – Andrew Murphy (PHE)
- maintenance manager – Andrew Murphy (PHE)

COSD has 2 main email addresses:

- COSD (main) – COSD@phe.gov.uk
- COSD (enquiries) – COSDenquiries@phe.gov.uk

[5] http://digital.nhs.uk/isce/publication/DCB0147
[6] http://digital.nhs.uk/isce/publication/SCCI0111
[7] http://digital.nhs.uk/isce/publication/DCB1533
[8] http://digital.nhs.uk/isce/publication/scci1577
[9] http://digital.nhs.uk/isce/publication/SCCI0021
[10] http://digital.nhs.uk/isce/publication/scci0034
[11] http://digital.nhs.uk/isce/publication/DCB2094
[12] http://www.rcpath.org/professional-standards

Health and care organisations

Requirements

NHS Providers of cancer services (hereinafter referred to as NHS Providers) must:

- read the Specification and Change Request in conjunction with the 'Implementation guidance' to identify how the standard is applicable to them
- review their system compatibility against this standard to identify any changes required to current practice to ensure that all data items in COSD v9.0 or COSD Pathology v4.0 can be flowed electronically by the dates specified in this document (if there are compatibility gaps, then further development is required) to meet the standard
- submit the data using the XML format for extracts from MDT cancer information management systems
- submit the data using the XML format for extracts from pathology systems

NHS Providers should not utilise this data set primarily to support their clinical and operational data capture.

It is important that where a Trust originally records a patient as having cancer and a record is sent during routine data uploads, but this diagnosis changes to a non-registerable condition, that NCRAS is immediately informed of this decision. Due to the complex way cancer information systems are designed, this change of status will not be sent automatically within the next available upload of data.

Conformance criteria

The following series of conformance requirements must be read by all health care organisations:

- data items submitted are as specified in the COSD v9.0 or COSD Pathology v4.0 data sets and submitted within the defined time period and in the format specified
- all specified linkage items are required (at record level) to enable linkage of the relevant cancer registration records
- there is a 25 working day reporting period following the month end, to submit the agreed data items following diagnosis date

- there is a 25 working day reporting period following the month end, to submit the agreed data items following treatment start date
- there is a 25 working day reporting period following the month end, to submit any additional or amendments to the data items
- an agreed method of submission with the NCRAS is required, for all items not flowed as part of the standard extract (for example, imaging data)
- notify the NCRAS (as soon as possible after discovery) of any known reasons for significant variation in the number of new cases submitted monthly in comparison to previous months
- monthly feedback from the NCRAS is provided for review, using the CancerStats portal[13]. This will allow cancer teams to assess if the data uploaded meets their expectation, and if not then they should be challenged with the Cancer Services Manager
- quarterly feedback reports are provided by the NCRAS to help the auditing of case ascertainment, quality and completeness
- NCRAS should be informed of reasons for any discrepancies as soon as possible
- A minimum of 80% of all expected cases is to be reported by provider Trusts annually by site specific tumour group, as agreed with the NCRAS
- all data extracted from Trust MDT cancer information management systems is to be reported in XML
- all data extracted from Trust pathology systems are to be reported in XML, or agreed method of reporting whilst Trusts migrate to new systems capable of reporting in XML

Conformance is measured against the COSD Conformance Framework which has been published on the COSD webpage[14]. Basic feedback on conformance is provided to Providers through the NCRAS, COSD CancerStats portal as reference below in footer[13].

Additional reporting for the National Prostate Cancer Audit[15] (NPCA), National Lung Cancer Audit[16] (NLCA), Incidence, Survival and Mortality data, are also available on the portal, along with Clinical Headline Indicators (CHI)[17].

Feedback reports are provided to the COSD Governance Board in order to monitor and manage compliance to the Information Standard.

---

[13] https://nww.cancerstats.nhs.uk/users/sign_in
[14] http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd
[15] https://www.npca.org.uk/
[16] https://www.rcplondon.ac.uk/projects/national-lung-cancer-audit
[17] www.ncin.org.uk/view?rid=2805

IT systems

It would be expected that all Trusts have a service level agreement (SLA), with their system supplier to ensure future development needs are sufficient to meet changes to the standard.

Requirements

Trusts must ensure that cancer IT and pathology systems change in accordance with their local contractual arrangements, to enable all specified data items in COSD v9.0 and Pathology v4.0 to be captured and extracted in compliance with the Specification and Implementation Guidance.

A new schema will be issued to reflect all changes for both COSD v9.0 and Pathology v4.0, which will support this process.

Conformance criteria

The above requirements must be met.

NHS Providers must submit the agreed data items within 25 working days of the month end following diagnosis or treatment date. The DCB and PHE provide all the documentation to support this process, including:

- implementation guide
- change request
- specification document
- COSD data set v9.0
- COSD pathology data set v4.0
- schema(s)
- data set user guides
- technical guides

Each regional NCRAS office has a nominated Cancer Improvement lead who will help and support any Trust struggling to meet the standard.

Scope

This is a change to the standard which introduces amendments to and re-aligning of data within the current data sets and a revision of the current schema specifications in order to continue to meet the business objectives of the standard (DCB1521 Amd 13/2019).

This version removes the ability to collect the pathology data outside of the pathology data set (v4.0) and is supported by the RC Path. This will reinforce the commitment to reduce the burden of data collection wherever possible and remove unnecessary duplication, and formally removes 155 data items from the COSD v9.0 data set.

As the pathology data items still exist albeit in the other existing pathology data set, the developer did not include these within the documentation as deleted items within COSD v9.0 as this may have looked disingenuous.

The data set relates to all cancer patients, both adult and paediatric, in acute inpatient and outpatient settings. The trigger for data collection is when a diagnosis, or suspected diagnosis of cancer is confirmed. Primarily this diagnosis takes place within secondary care.

The standard covers neoplasms coded within ICD-10 diagnosis codes range C00 - C97, D00 - D48 and E85.918. See Appendix B for list of Mandatory Registerable Conditions.

All changes can be found within the published Change Request document.

Wherever possible duplication across the data set has now been removed and full explanations of how to collect these data within the new structure are provided within the change logs of each data set.

Finally, where there were data that are no longer part of a linked national data set (for example, the Royal College of Pathologists), these have also been removed from v4.0 of the COSD Pathology data set.

---

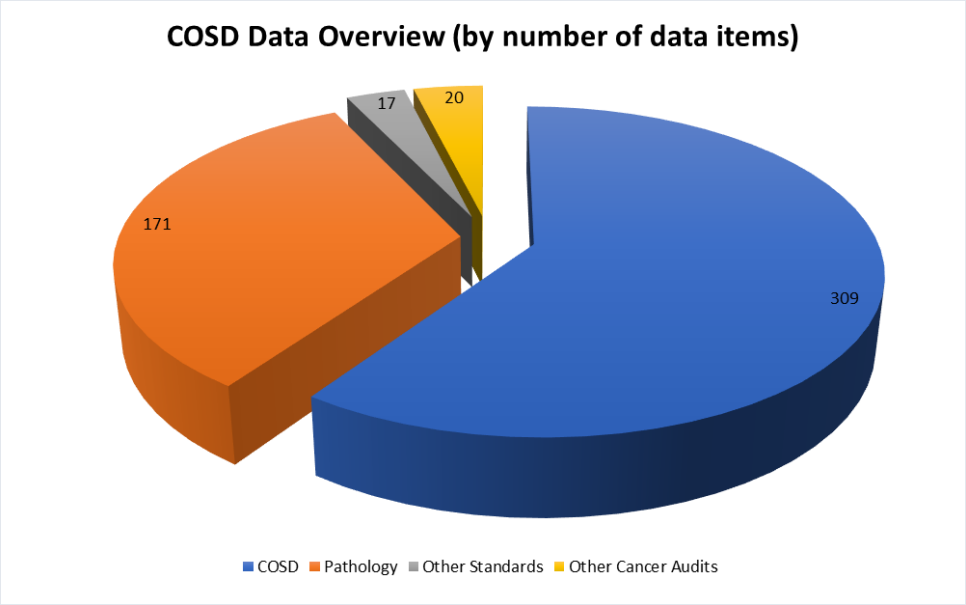
[18] Although Primary amyloidosis (E85.9) is listed as an E ICD code in the World Health Organisation (WHO) disease classification, amongst clinicians it is widely acknowledged and subsequently treated as a cancer, receiving Chemotherapy in cases. Whilst we await the WHO disease classification being updated to reflect this fact, we have extended the scope of the COSD to include this. The United Kingdom and Ireland Association of Cancer Registries (UKIACR) is currently considering its inclusion in the UKIACR Library of Recommendations, which we have referenced in Appendix B.

Implementation and use

Guidance

This standard defines the complete set of secondary uses cancer data for reporting and specifies the items which need to be returned directly by NHS Providers. Fig 1 below helps understand the responsibility for collecting/reporting data.

Fig 1: COSD compilation overview



Provider submissions

The 309 COSD items shown above comprise the subset of the COSD to which the remainder of this specification refers unless otherwise stated. These are the items which are included in the XML schema and are expected to be flowed directly from NHS Providers to the local NCRAS Branch Office, from one or more electronic systems within the Provider organisation.

Pathology

The 171 pathology items are expected to be collected directly from the pathology department and are identified separately within their own data set. These are items which are essential to compile the full data set but are covered by their own schema and data flow.

Other data

Other items are either subject to ‘Other Standards’ (such as CWT, SNOMED CT or Person Sexual Orientation), or are collected primarily for ‘Other Cancer Audits’ but are requested as additional supporting data through COSD. This excludes cancer audits directly collected through COSD e.g. Lung, Breast and Prostate. Further details are provided in user guidance, and wherever possible duplication has been removed, and in most cases these data are all collected within the same collection/reporting system.

Model data flow diagram

The following diagram (fig 2) demonstrates how the full COSD data set will be compiled centrally by local NCRAS offices from data flowing from a number of systems and sources.

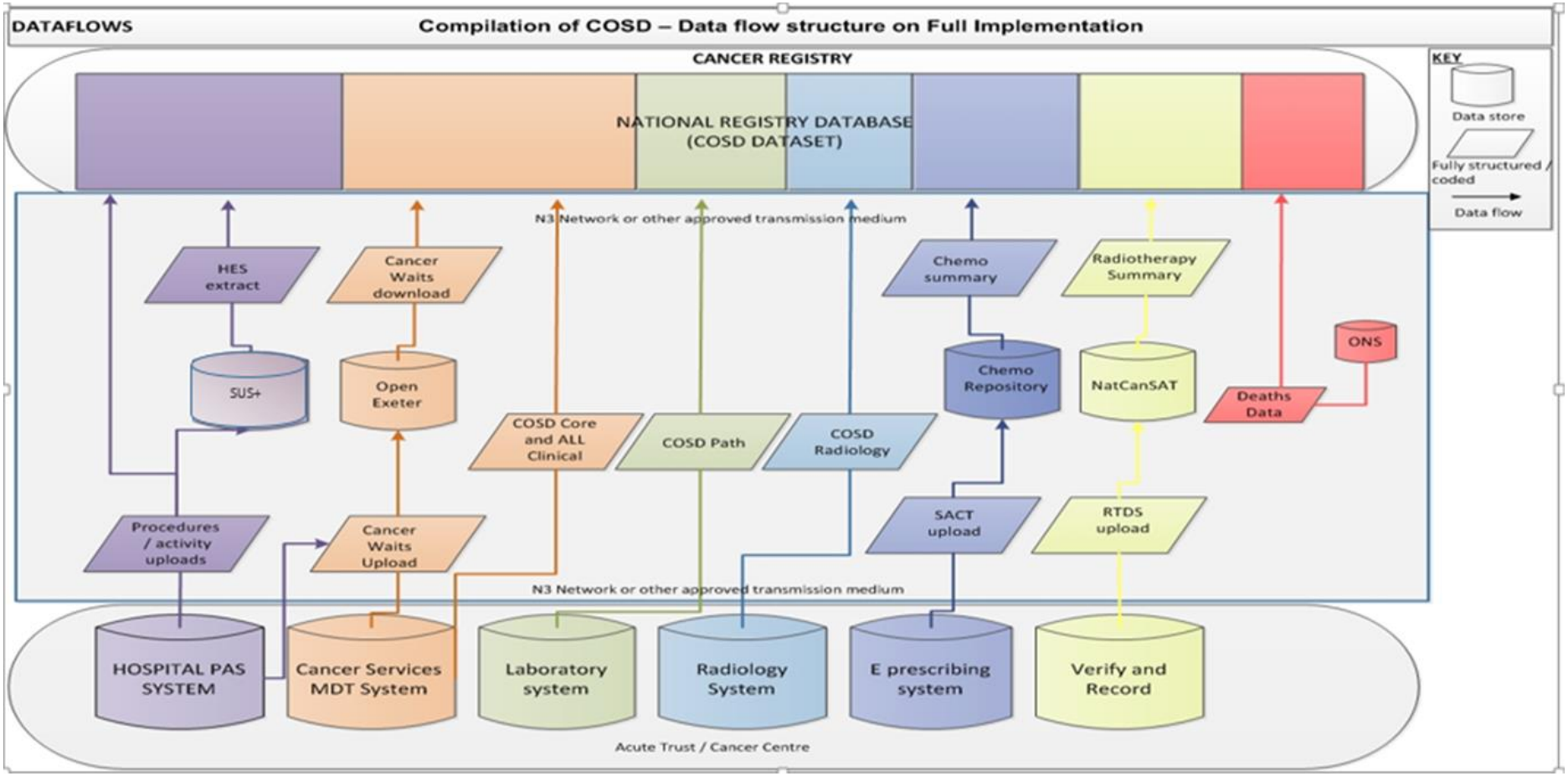


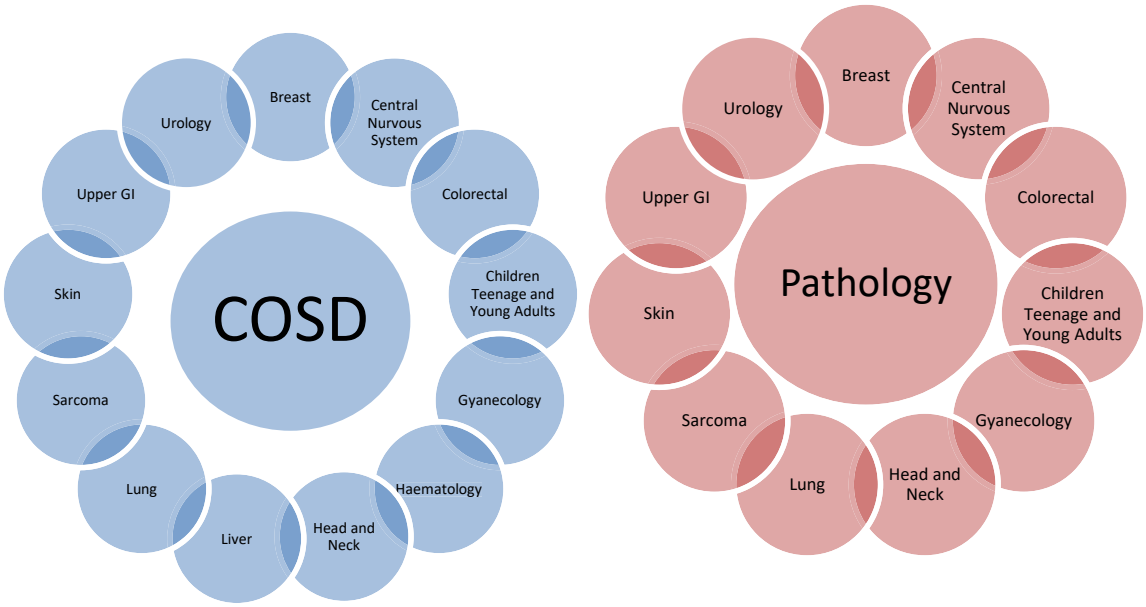
Fig 2: COSD compilation - Data flow structure on full implementation

Structure of data set – provider submissions

There is a core data set, of which most data items are applicable to all cancers recorded, and an additional set of site specific data sets (one for each of the thirteen identified tumour groups). Some of these site specific data sets contain further subsets applicable to individual diseases.

Each recorded case will therefore have a core and usually a site specific data set completed (see fig 3 below). This would be the same for pathology, where you have a Core and then eleven site specific pathology data sets.

Fig 3: Structure of COSD



Data set subsections

Within each of the core and site specific data sets, the data items are further grouped according to their stage along the patient pathway. In both data sets, choices have been introduced to define decisions better and improve data quality. Several data items have also become mandatory from 2020, in an effort to improve data quality and prevent records being submitted without enough information to be useful for cancer registration.

The user guides (available as separate documents[19]), provide clear explanation's on data flows for a primary cancer and non-primary cancer pathway. This includes explanations on what is a recurrence, transformation and progression and supporting flow diagrams.

[19] http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd

Governance

Information governance, clinical safety and data protection

The primary purpose of the standard is for secondary uses only and will therefore have no direct impact on clinical safety and as such is not in scope of DCB0129[20]. Consequently, a clinical safety case report is not required to support the standard.

A full data protection impact assessment has been completed by PHE and submitted to the DCB as part of the supporting evidence for this updated standard.

NCRAS is part of PHE which is an executive agency of the DHSC. The function of PHE is to fulfil the Secretary of State for Health's statutory responsibilities to protect and improve public health and reduce health inequalities. Running national data collections on a range of diseases, including cancer, is a vital part of this work.

Section 251 of the NHS Act 2006 provides the statutory power to ensure that NHS patient identifiable information needed to support essential NHS activity can be used without individual patient consent. This power can be used only to support medical purposes and the interests of the public if seeking consent is not practicable and the use of anonymised information is not sufficient.

Under Regulation 2 of the Health Service Control of Patient Information Regulations 2002, NCRAS is authorised to process: "confidential patient information relating to patients referred for the diagnosis or treatment of neoplasia."

Permission to process these data is subject to annual review by the Confidentiality Advisory Group (CAG) of the Health Research Authority. The registry's performance against national Information Governance (IG) standards for health data is assessed annually on the NHS Data Security and Protection Toolkit.

The NCRAS must achieve at least a Level 2 score each year to continue to process patient identifiable information under Section 251. CAG advises the Secretary of State on whether to grant approval to continue to process based on its review and the toolkit score.

[20] DCB0129 Clinical Risk Management: its Application in the Manufacture of Health IT Systems: http://digital.nhs.uk/isce/publication/dcb0129

Under General Data Protection Regulation (GDPR) the lawful basis upon which the registry will process personal data is Article 6(1) (e): “processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority.”

The registry receives health and genetics data in accordance with the conditions for “special category” data set out in GDPR Article 9(1) (h): “processing is necessary for the… provision of health care or treatment or the management of health...care systems and services.”

And GDPR Article 9(2)(i): “processing is necessary for reasons of public interest in the area of public health such as... ensuring high standards of quality and safety of health care… on the basis of [UK] law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy.”

Reported data will be managed by the NCRAS where there is long standing expertise in managing large volumes of confidential data. Although the data items which are flowed to the NCRAS have changed, the data flows (such as which organisations will be receiving the data in identifiable form) remain unchanged.

In compliance with the fair processing requirement within the Data Protection Act, provider organisations are expected to inform patients of this purpose for reporting their information and of the potential use of the information for service development, analysis and statistical research.

NCRAS supports this process and provides privacy information in several ways. Its patient leaflet, for example:

- explains what cancer registration is, why it’s important, where to go to find more information and how to opt out
- was designed in partnership with patient groups and cancer charities and was approved by the Plain English Campaign
- with 196,000 of them distributed in the first 6 months of 2018 to all NHS Acute Trusts in England
- which were sent to Trusts’ cancer services, patient information centres, many site-specific Multi-Disciplinary Teams (MDTs), cancer charities and private health care providers
- which the NCRAS includes in patient surveys
- which is widely available across the internet[21]

Alongside the leaflet NCRAS has also distributed around 500 posters to put up in outpatient clinics, information centres and waiting areas

[21] https://www.ndrs.nhs.uk/wp-content/uploads/2019/01/Cancer-Registration-information-leaflet-JAN-19-WEB.pdf

NCRAS, as part of PHE, complies with the DHSC's Data Protection Act registration with the Information Commissioner's Office (ICO). The NCRAS regularly reviews and harmonises its information governance policies to correlate them with those of PHE and aligning with multiple mandatory training requirements (annually) for its employees.

These policies inform, for example:

- access controls of data
- server security and encryption
- data transfer procedures

All NCRAS employees handling patient identifiable data (PID), are required to complete information governance, data security and responsible for information mandatory training. There is also a 'Confidentiality Guidelines and Agreement' document, which is an individual declaration for all employees and must be read and signed annually. This is monitored as part of its appraisal process.

Consent process

Where patients have requested their data are not shared, the provider organisation must ensure that their records are not included in the data downloads submitted to the NCRAS. It is suggested that a dissent (such as the proactive expression by an individual from whom consent has not been obtained) or a similar flag should be provided in the provider organisation systems so that the record can then be omitted from the monthly upload.

The NCRAS has published a Patient Information Leaflet (Appendix C) which explains that individuals have the right to access and have their own data held in the NCRAS removed and explains the process.

If a patient discovers that their information has been uploaded to the NCRAS and they wish for this to be deleted, the requester can email their request to: optout@phe.gov.uk, or write to the Director of NDR[22] using the address in the patient leaflet. The NCRAS will then, as far as is possible, remove the patient from its database.

The NCRAS information for patients wishing to have their information removed as far as possible from the NCRAS database is available on the NCRAS website (http://www.ncras.nhs.uk/patientinfo).

[22] National Disease Registration

Data retention

The NCRAS holds data indefinitely, as without this facility accurate mortality and survival data cannot be calculated. If a patient wants their data removed at any point, then this facility is available as explained above.

Data disclosure

The NCRAS adheres to the requirements of the Data Protection Act 1998 with regard to the receipt, storage and transfer of information relating to individuals. When releasing data to third parties, all NCRAS offices strictly comply with the approval on the release of patient-identifiable and potentially identifiable information.

Recipients of such data are required to sign a declaration stating that they will protect the information they are entrusted with, use it only for the purpose for which it was supplied and make no attempt to identify information pertaining to particular individuals or to contact individuals. They are also prohibited from presenting any information that may identify an individual. This is also the case with publications produced by NCRAS, which present aggregated data only.

As the NCRAS is part of PHE, all such requests must be approved by the Office of Data Release (ODR).

Subject access requests

Subject Access Requests (SARs) are managed through PHE's Public Accountability Unit, in line with PHE's Data Protection Policy and the Data Protection Act.

For SARs, relating to cancer registration, there is a 2 stage information release process. This is designed to balance the needs of the individual with our statutory requirement to protect patient confidentiality.

At the first stage, we provide the subject with a summary report known as an 'exemplar report'. This report contains information about data sources attached to the subject's cancer registration, the date we received and the originating organisation. The report gives examples of the type of information these sources might contain and a glossary of terms.

Following this, If the subject requires more information, we also offer to send a copy of the actual records attached to the subject's cancer registration to the primary care physician (the General Practitioner) or another clinician who knows the patient.

This allows the GP to verify that the data has been sent to the correct person and, because some of the information is quite technical, the recipient also has the opportunity to ask the practice staff to explain anything they may not understand.

All the information is reviewed by the NCRAS' own Caldicott Guardian before being released. Requests are fulfilled within the time periods required by the Data Protection Act.

There is a sample process map attached in Appendix D.

Data quality

Two areas for consideration are the quality of data submitted by providers and the data quality processes at the NCRAS offices.

Before and during the review process for COSD v9, previous versions were reviewed in detail and the submitted data assessed to ensure that it was still fit for purpose. In addition, it is important that the data being received continues to be of a standard that can influence national analysis and international benchmarking.

This review process is important as it identifies data that requires clinical review and if necessary either changing, updating or removing from future versions. This can happen due to changes to clinical practice and also the importance of the data for analytical reporting. Priorities change as does the direction of medicine over time, and COSD has to reflect this.

This constant review is an important part of the development process. As a result of this review, several data items have now become mandatory across the data sets, enforcing stronger data quality, validation and improving ascertainment.

Equally many data items have also been removed as they were difficult to collect, or the quality was so poor it was deemed not practicable to continue requesting them as they were not influencing outcomes or analysis but were impacting of the overall burden of data collection.

NHS Providers

Each Provider is responsible for ensuring the data submitted to the NCRAS or submitted through other standard NHS routes is of the highest quality and completeness possible, and accurately represents the service provided.

The NCRAS provides a dynamic feedback process from the cancer registration system to Providers. This will allow data quality assurance at a field level - with clinical teams given secure access to the data (via the Cancer Stats portal) that their organisation has submitted.

National Cancer Registration and Analysis Service

One of the main roles of the NCRAS is to ensure data quality and consistency. The 8 NCRAS regional offices use a single online processing system called (ENCORE). Working practices have been standardised with continuous performance monitoring and oversight of the entire NCRAS through Public Health England.

In 2016, NCIN became part of the NCRAS, which is part of PHE. This has enabled more efficient analysis of cancer data. Specific aspects of data quality are described in Appendix E.

CancerData[23] is a publicly accessible portal to look at outcomes data for CCG and provider Trusts. This has been specifically released for the public as well as NHS organisations, and is a vital step forward in improving data quality, whilst producing as near-to real-time data analysis from the NCRAS.

Although these data are not drawn exclusively from COSD 'level 2' measure this in CancerStats. There are 3 levels within CancerStats for COSD reporting:

- level 1 – which measures submissions criteria
- level 2 – which measure performance of certain data exclusively from the Trust
- level 3 – which looks at all data reported to and processed on a patient's cancer registration record, regardless where the data came from

COSD does form part of the total data used within the NCRAS cancer registration record and is an important part of data collection across the NHS in England.

[23] https://www.cancerdata.nhs.uk/index.html

Demographic data

The cancer registration data is dynamic and individual tumour records are updated from numerous disparate data sources. Linkage of some of these sources across the NHS is not sufficiently good to allow accurate mapping of new data to existing items without patient-identifiable data.

An example could be when using only NHS Number. There is a risk that a typo could cause the wrong data to be inadvertently added to the wrong patient or a new record created to a patient who does not have cancer. Having additional patient identifiers helps to remove that possibility and provides an invaluable quality assurance process between the Trusts and the NCRAS.

Even once linked, retaining addresses and names remains important; the address stored by the NCRAS is that at the time of diagnosis of the tumour and is essential for cancer cluster analyses possibly many years later, when the patient may have moved. Without patient name, registries could not support genetic and follow-up enquiries made by clinicians who often only have limited information on the index case and possible relatives.

Nevertheless, it is registry practice to use pseudonymised or even anonymous (possibly still disclosive) data sets for analysis where patient identity is not needed. Access to identifiable and potentially disclosive data requires appropriate permissions from the Office of Data Release (ODR). This is a function of the National Disease Registry (NDR) within PHE.

For details on how the cancer registry processes deal with linkage and data discrepancies, please see Appendix F.

Contact details

Information, including the COSD data sets and COSD user guides are available on the NCIN website at:
http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd

Queries regarding this document should be addressed to:
COSDenquiries@phe.gov.uk

Queries regarding submissions should be discussed with the NCRAS regional liaison managers. Contact emails and telephone numbers are available here:
http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd

Appendix A: glossary of terms

The following is a glossary of terms which relate to and are referenced within this specification document. The term is followed (in brackets) by the acronym, then below is the definition:

Burden

The Data Standard Assurance Service (DSAS) team within NHS Digital assess the burden on the NHS of all data collections as per the Health and Social Care Act 2012. Anyone who collects data on a national scale from the service needs to go through the Burden process to gain approval to collect the data.

Cancer

For the purposes of this standard the term ‘cancer’ is used throughout the standard and related documents to cover all conditions defined by the World Health Organization (WHO) and International Agency for Research on Cancer (IARC) Classification of Tumours[24].

Cancer centres

Organisations which help people to live with, through and beyond cancer by bringing together specialist clinical and professional staff and communities of support.

Cancer Outcomes and Services Data set (COSD)

The COSD is the national standard for reporting cancer in the NHS in England. It replaced the previous National Cancer Data set and includes the former Cancer Registration data set and additional site specific data items relevant to the different tumour types.

Cancer Registries

Organisations which exist internationally to collect, process, analyse and disseminate data on cancer patients in their local regions.

[24] http://whobluebooks.iarc.fr/

Cancer stats portal

Online resource that allows Trusts to review specific cancer data submitted as part of the conformance framework.

Care Quality Commission (CQC)

One of the independent regulators of health and social care in England[25].

Commissioners

Organisations that plan, purchase and monitor services to meet the health needs of their local population.

Department of Health and Social Care (DHSC)

DHSC is a ministerial department, supported by 28 agencies and public bodies[26].

Diagnostic Imaging Data set (DIDS)

Data set containing diagnostic imaging test activity across the NHS, taken from Radiology Information Systems. (See SCCI Standard 1577)[27].

Expert Advisory Groups (EAGs)

A group of experts (at tumour site level) who advise NCRAS on what data needs to be collected and what analyses conducted.

Extensible Markup Language (XML)

Extensible Markup Language (XML) is a set of rules for encoding documents in machine-readable form.

Information Standard Notice (ISN)

Information Standards Notices (ISNs) are published to announce new or changes to information standards published under section 250 of the Health and Social Care Act 2012.

---

[25] https://www.qcs.co.uk/
[26] https://www.gov.uk/government/organisations#department-of-health-and-social-care
[27] http://digital.nhs.uk/isce/publication/scci1577

International Statistical Classification of Diseases and Related Health Problems (ICD)

A medical classification list for the coding of diseases, signs and symptoms, abnormal findings, complaints, social circumstances and external causes of injury or diseases, as maintained by the World Health Organization (WHO). The title is followed by the revision number, e.g. ICD10 is the tenth revision.

International Classification of Diseases for Oncology (ICD-O)

An extension of the ICD coding system used principally in tumour or cancer registries for coding site (topography) and histology (morphology) of neoplasms. The title is followed by the revision number, e.g. ICD-O-3 is the third revision.

Multi-Disciplinary Team (MDT) coordinator

The MDT coordinator is the person(s) responsible for facilitating the MDT meeting. They also have additional duties for collecting and recording information on patients as they pass through the Provider Trust, whilst on a cancer pathway. Sometimes known as a Patient Pathway Coordinator.

National Cancer Data set (NCDS)

The previous nationally approved reference standard for the collection of cancer data now incorporated into the COSD.

National Cancer Registration and Analysis Service (NCRAS)

The NCRAS is the National Cancer Registration and Analysis Service for England, collecting cancer data from all NHS Providers of cancer care in England. The NCRAS is a function within the National Disease Registration Service within PHE.

National Cancer Waiting Times Monitoring Data set (NCWTMDS)

The Information Standard (DCB0147)[28] used to monitor the time that patients with suspected and diagnosed cancer have wait for appointments, tests and treatments.

[28] http://digital.nhs.uk/isce/publication/SCCI0147

NHS Digital (NHSD)

NHS Digital[29] is England's central, authoritative source of health and social care information for frontline decision makers, which builds upon the Health and Social Care Act 2012.

NHS England (NHSE)

NHS England is an executive non-departmental public body (NDPB) of the Department of Health and Social Care. NHS England[30] oversees the budget, planning, delivery and day-to-day operation of the commissioning side of the NHS in England as set out in the Health and Social Care Act 2012.

NHS Improvement (NHSI)

From 1 April 2019, NHS England and NHS Improvement[31] are working together as a new single organisation to better support the NHS to deliver improved care for patients.

Office of Data Release (ODR)

The ODR was established by the senior information risk owner to provide a systematic, risk-based approach to reviewing both internal and external requests to process PHE data for secondary purposes and to meet our obligations in line with data protection legislation, Common Law Duty of Confidentiality, Caldicott Principles and best practice set by the Information Commissioner’s Office (ICO).

Office of National Statistics (ONS)

The UK's largest independent producer of official statistics and the recognised national statistical institute of the UK.

Providers

Organisations that provide health services.

---

[29] https://digital.nhs.uk/
[30] https://www.england.nhs.uk/
[31] https://improvement.nhs.uk/home/

Public Health England (PHE)

Public Health England[32] is an executive agency of the DHSC. Its role is protecting and improving the nation's health and wellbeing and to reduce inequalities.

Radiotherapy Data Set (RTDS)

A standard data set covering every patient treated with radiotherapy in the NHS in England[33].

Systemic Anti-Cancer Therapy Data Set (SACT)

The national collection of all cancer systemic anti-cancer therapy data in the NHS in England, which covers all solid tumour and haematological malignancies. This includes all adult and paediatric cancer patients, those in clinical trials, and covers acute inpatient, day case, outpatient and community settings[34].

The Royal College of Pathologists (RC Path)

A professional membership organisation committed to setting and maintaining professional standards and to promoting excellence in the practice of pathology.

United Kingdom and Ireland Association of Cancer Registries (UKAICR)

The UKIACR brings together organisations with an interest in developing cancer registration as a resource for studying and controlling cancer in the UK and Ireland[35].

XML schema

The documentation, definitions and descriptions required to enable the production and transmission of data for a specific XML.

---

[32] http://phenet.phe.gov.uk/Pages/Home.aspx
[33] http://digital.nhs.uk/isce/publication/SCCI0111
[34] http://webarchive.nationalarchives.gov.uk/+/http://www.isb.nhs.uk/documents/isb-1533/amd-24-2013
[35] http://www.ukiacr.org/

Appendix B: mandatory registerable conditions

ICD 10	Description of neoplasm
C00 - C97	All malignant neoplasms
D00 - D09 (excluding D04)	All carcinoma in-situ (excluding all D04 in-situ skin cancers)
D32 - D33 D35.2 & D35.3 D35.4	Benign neoplasms of brain & other parts of nervous system Benign neoplasms of pituitary gland & craniopharyngeal duct Benign neoplasms of pineal gland
D37 - D48 (excluding D47.2)	All neoplasms of uncertain behaviour Neoplasms of unspecified nature of bladder Neoplasm of unspecified nature of brain Neoplasm of unspecified nature of other parts of nervous system & pituitary gland only (Excluding D47.2 Monoclonal gammopathy of undetermined significance (MGUS))
E85.9	Primary Amyloidosis[36]

Please see COSD User Guide for full list of Mandatory Registerable Conditions.

[36] Although Primary amyloidosis (E85.9) is listed as an E ICD code in the World Health Organisation (WHO) disease classification, amongst clinicians it is widely acknowledged and subsequently treated as a cancer, receiving Chemotherapy in cases.

Appendix C: cancer registration Leaflet

Below is the cancer registration leaflet (v6), as of 6 January 2019.

Where can I find out more?

If you would like to find out why cancer registration is important or have any questions about the work we do, you can:

- visit us online at www.ndrs.nhs.uk
- talk to a member of the NHS cancer team treating you, or
- visit www.nhs.uk/your-data-matters to find out how the NHS uses information.

Public Health England

Protecting and improving the nation's health

NHS

Can I see the information you hold about me?

Yes, we can give it to a doctor (GP) who knows who you are, so they can share all the information with you.

Can I ask for my information not to be included in the cancer registry?

Yes, you have the right to opt out of cancer registration. This will not affect the personal care you receive from your healthcare team.

If you do not want your information included in the national cancer registry, you can contact us at optout@phe.gov.uk or write to:

Director
National Cancer Registry
Public Health England
6th Floor, Wellington House
133-155 Waterloo Road
London SE1 8UG.

For information on your rights and privacy visit www.ndrs.nhs.uk/cancer-registration-your-rights-and-privacy

This leaflet is available in alternative formats. Contact us at NCRASfeedback@phe.gov.uk for more information.

This leaflet is reviewed regularly. If you have any comments, please email NCRASfeedback@phe.gov.uk.

PHE publications gateway number: 2018747. Version 6, January 2019.

Crystal Mark 22491 Clarity approved by Plain English Campaign

Cancer registration

Why it matters and what you need to know

What is cancer registration?

If you are diagnosed with cancer or a condition that may lead to cancer, the NHS team looking after you will record information about you and the care you receive. This applies to children and adults of all ages.

This information is shared with the National Cancer Registry, which is part of Public Health England.

The National Cancer Registry has the government's permission to collect and use information about people with cancer. This is because it is in the public interest to use this information to improve the way cancer is diagnosed and treated.

Why it matters

Cancer registration is the only way we can know how many people are getting cancer and the types of cancer they have.

This information helps us to:

- look at overall trends in cancer
- improve the diagnosis of cancer
- develop new treatments and drugs
- improve cancer services, and
- inform national cancer policy.

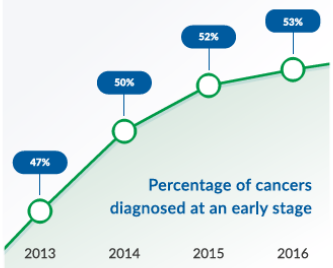
What information is collected?

The information we collect includes:

- your name and date of birth
- your sex and ethnic background
- your address and NHS number
- information about your diagnosis, and
- information about your treatment and how well your treatment is working.

This information may be linked to other health information that we may receive about you.

It is really important that cancer is diagnosed as early as possible. Cancer registration supports the work to improve earlier diagnosis.



The diagram shows there has been an increase in the percentage of cancer cases that are being diagnosed at an early stage, when treatment is more likely to be successful.

How will it benefit me?

We know that cancer registration is leading to improvements in preventing, diagnosing and treating cancer. This benefits everyone affected by cancer.

Healthcare staff may use information from the cancer registry to see if you might benefit from being part of a clinical trial.

Some cancers run in families. With your permission, doctors can use your information to see if other members of your family may be at risk, and find the best ways to treat them.

Is my information secure?

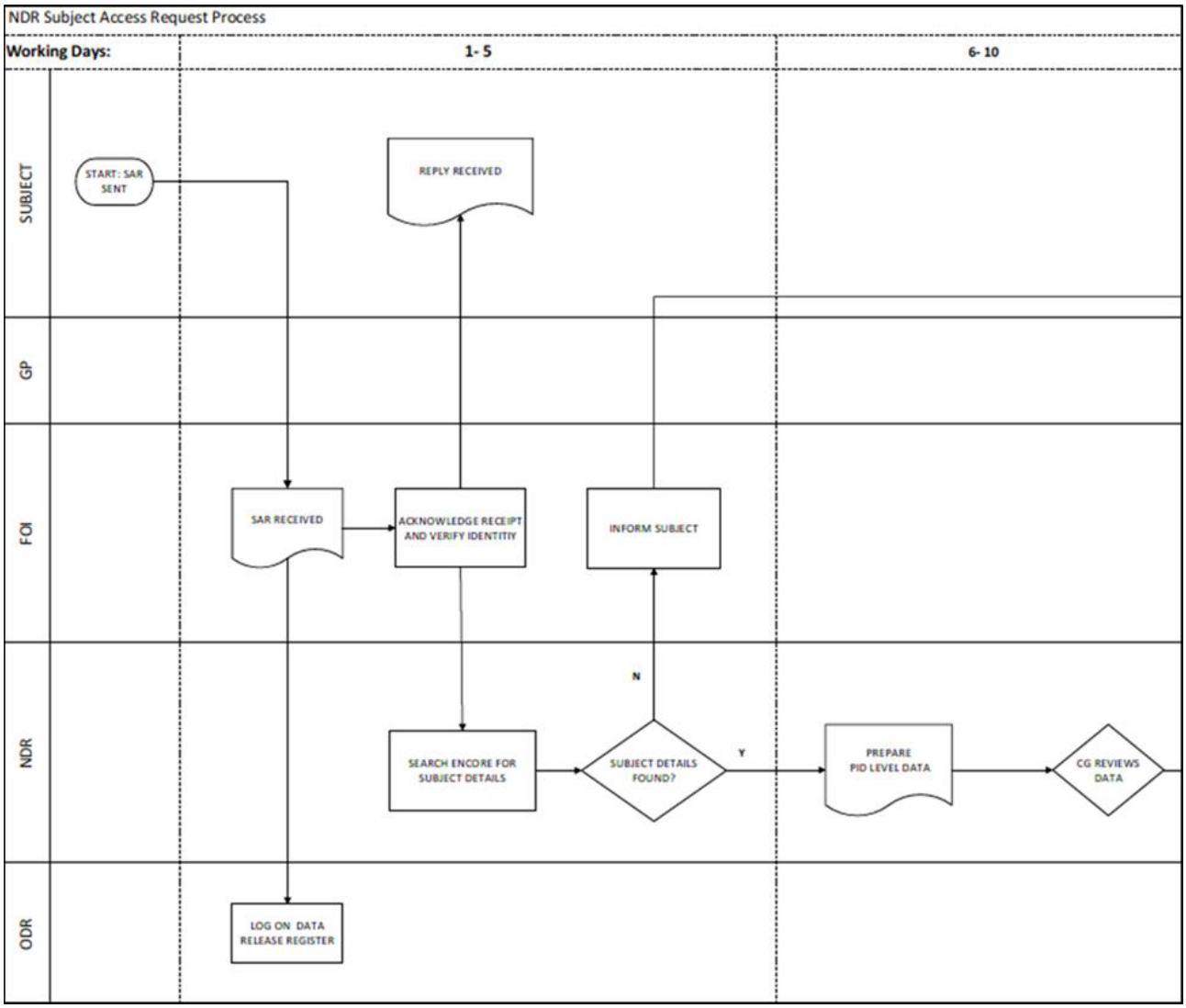
Yes, all your information is kept confidential.

Cancer registration helps drive research into cancer so we may sometimes need to share your information with researchers outside Public Health England. There are very strict rules for this. It only happens if the researchers have a lawful reason to use the information. Researchers must prove that the information will be kept safe and secure to protect your privacy.

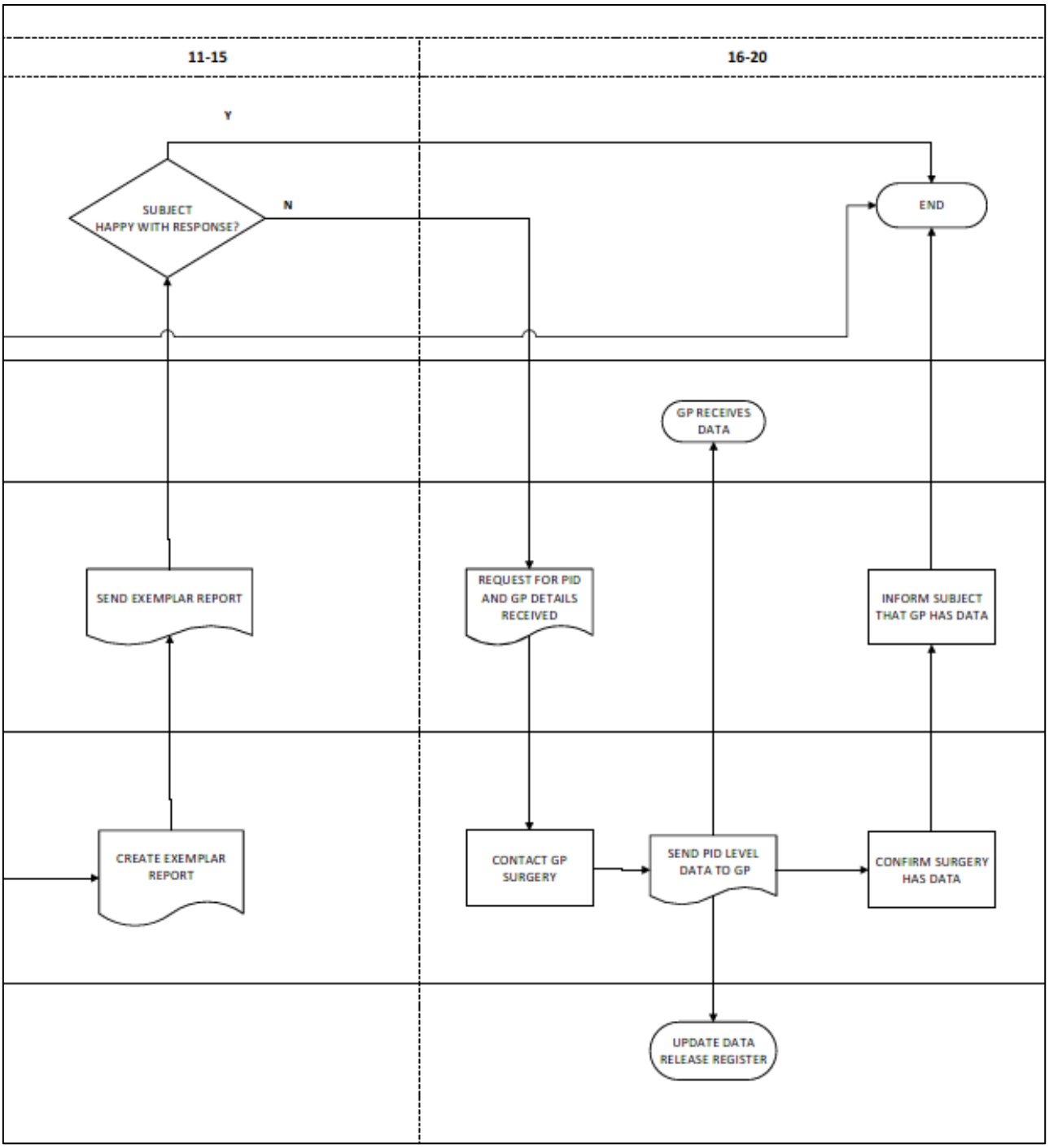
Appendix D: subject access request

The 2 following flow diagrams describe the subject access request process.

FOI days 1 to 5 and 6 to 10



FOI days 11 to 15 and 16 to 20



Appendix E: NCRAS data quality controls

Automated Quality Control

The data files submitted through the NCRAS data clearing house are subject to a wide range of validation rules to ensure that the data files and data within fields is consistent, as explained in the 5 sections below:

Batch tracing of all cases

All patient-identifiable electronic records are sent to the Demographics Batch Service for tracing against the NHS Spine, where discrepancies are identified, investigated and whenever possible reconciled.

This will not place any additional load on either the Personal Demographic Service (PDS) or Batch Tracing facilities, however the impact will be monitored by the NCRAS as the project progresses with any significant increase being brought to the attention of the COSD Governance Board.

Use of multiple data sources

The quality of cancer registration relies upon the use of multiple independent data sources to ensure high ascertainment and cross validation, the ENCORE system automates much of the data linkage between the disparate sources - highlighting inconsistencies that can be further investigated.

Cancer registration staff

The National Cancer Registration and Analysis Service employs tumour registration staff at the local registries. These registration staff have considerable expertise in cancer coding and classification and spend much of their time quality assuring the electronic data sources and cases recorded at the registry. In some cases, cancer registration officers work in a local provider organisation, but all registries maintain very close contacts with the clinical teams.

Data feedback to clinical teams

Rapid feedback to the provider clinical teams, usually through the MDT provides an important process of data validation, the NCRAS uses secure web-based systems to deliver reports at a field-level on the completeness of individual data items.

Data quality audit

The UKIACR has developed a large number of performance metrics covering the process of data collection by registries, these performance metrics have been integrated into the new ENCORE system and will (where appropriate), form the basis of daily updates on the data quality and completeness of records held in the NCRAS.

Appendix F: data linkage and data discrepancies

Linkage

Linkage is a complex issue, which has become far simpler in recent years with the rollout in use of the NHS Number. Registries use different linkage methods according to the type of data which is available. In essence, the more data that is available, the more confident that linkage is correct. In fact, linkage comprises 2 parts; blocking and weighting.

Blocking

This takes an incoming record and uses a range of search criteria, determined by the incoming records content, to identify a series of possible matches in the database:

- where the NHS Number is available that is used, but other blocking is usually also applied
- in a manual context, these blocks tend to be sequential, but in an automated setting they tend to run consecutively, with all potential matches passing to the second stage, weighting

Weighting

Weighting can be simple, including:

- deterministic weighting which is used for NHS Number matching – but this is always augmented with at least one other identifier
- probabilistic techniques use a wider set of data matches - and are usually used when the NHS Number is not available on either the source record or the blocked record
- looking for the ‘commonness’ of the data value in the overall database, and then uses that to weight up or down based on a series of random control matches
- probabilistic weighting is a well-defined science, with robust methodologies, however it is used far less than in past years

Data discrepancies

The fundamental principle of cancer registration is that it relies on multiple sources of data. When dealing with multiple sources, many of which may contain a common item, there is likelihood that 2 sources will give different values for an item of data.

The technical design of the registration schema is such that multiple sources and multiple data values are held against the summarised registration record. Registration officers are trained to identify and deal with data discrepancy. This usually starts with some basic data checking with the source data supplier, but where conflicting data exists there are clear rules by which registry staff undertake this.

At no stage is any source data overwritten or lost, and regular checks are included in the registration practice to examine random sets of records as part of the standard QA built into registration practice over many years.