Document filename:	**Annex A: NCRAS Requirements Specification for the National Disease Registries Directions**		
Project / Programme	**NDRS**	Project NDRS	
Document Reference			
Project Manager		Status	**Published**
Owner	Sarah Stevens	Version	2
Author		Version issue date	**02/02/2023**

Annex A: Requirements Specification for the National Disease Registries Directions 2021:

National Cancer Registration and Analysis Service

Document management

Revision History

Version	Date	Summary of Changes
0.1	07/06/2021	First draft
0.2	06/07/2021	Uplifted following review
0.3	20/07/2021	Uplifted following review by NDRS colleagues
0.4	27/07/2021	NHSD Legal review undertaken
0.5	29/07/2021	Uplifted following NHSD Legal review
0.6	05/08/2021	Uplifted following review by Executive Director of PTE
0.7	16/08/2021	Specification split into two - this section addressing NCRAS
0.8	07/09/2021	Uplifted following internal review
0.9	07/09/2021	Document prepared for NDRS review
0.10	13/09/2021	Uplifted following review by NDRS colleagues
0.11	23/09/2021	Uplifted following EMT review
1.0	01/10/2021	Final version 1.0
1.1	12/01/2022	Uplifted to include of collection of SARS-COV-2 subset of SGSS
1.2	02/02/2023	DRAFT - Uplift to include collection of NICOR dataset & transplant data and removal of dissemination via ODR
2.0	02/02/2023	Final Version 2.0

Reviewers

This document must be reviewed by the following people:

Reviewer name	Title / Responsibility	Date	Version
Magi Nwolie	Associate DPO	07/06/2021	0.1
		20/07/2021	0.3
		29/07/2021	0.5
		05/08/2021	0.6
		17/08/2021	0.7
		07/09/2021	0.8
		13/09/2021	0.10
		23/09/2021	0.11
		12/01/2022	1.1
Lucy Elliss-Brookes	Associate Director for Data Curation	06/07/2021	0.2
		20/07/2021	0.3
		06/09/2021	0.7
		07/09/2021	0.9
Paul Davies	NDRS	20/07/2021	0.3

		13/09/2021	0.9
Jackie Gray	Executive Director of Privacy, Transparency and Ethics	04/08/2021	0.5
Mike Eden	NDRS Associate Caldicott Guardian	07/09/2021	0.9
Louise Whitworth-Woodhead	Associate Director of Operational IG	03/03/2023	1.2

Approved by

This document must be approved by the following people:

Name	Signature	Title	Date	Version
Sarah Stevens		National Disease Registration Service	02/02/2023	1.2

Glossary of Terms

Term / Abbreviation	What it stands for
NCARDRS	National Congenital Anomaly and Rare Disease Registration Service - a service provided by NHS Digital under the National Disease Registries Directions 2021
NCRAS	National Cancer Registration and Analysis Service - a service provided by NHS Digital under the National Disease Registries Directions 2021
NDRS	National Disease Registration Service - a programme team within NHS Digital

Document Control:

The controlled copy of this document is maintained in the NHS Digital corporate network. Any copies of this document 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.

Contents

1. Purpose of document

This document sets out the Requirements Specification for the National Cancer Registration and Analysis Service (**NCRAS**) and should be read alongside the:

- National Disease Registries Directions 2021 issued by the Secretary of State for Health and Social Care; and
- Requirements Specification for the National Congenital Anomaly and Rare Disease Registration Service (**NCARDRS**).

2. Introduction

This document sets out the Requirements Specification in relation to NHS Digital collection and analysis of information previously collected by Public Health England (**PHE**) about:

- Individuals in England diagnosed with cancer, treated for cancer, suspected of having cancer, or with certain conditions that may lead to cancer.

The data was previously processed by PHE in discharge of the Secretary of State's public health functions under section 2B and under section 251 of the National Health Service Act 2006 (**Section 251**).

- PHE processed the cancer data through the NCRAS. The data was processed without patient consent under Section 251 (ref: PIAG 03(a)/2001) and regulation 2 of the Health Service (Control of Patient Information) Regulations 2002 (**COPI**) (Medical purposes related to the diagnosis or treatment of neoplasia).

The National Disease Registration Service (which comprises the National Cancer Registration and Analysis Service and the National Congenital Anomaly and Rare Disease Registration Service) has been transferred to NHS Digital under the National Disease Registries Directions 2021 (**Directions**). Under the Directions, NHS Digital is directed under:

- Section 254(1) and (6) of the Health and Social Care Act 2012 (**the 2012 Act**) - to establish and operate the National Disease Registries Information System (the **Information System**) for the collection and analysis of information as described in this Requirements Specification and in the related Requirements Specification for the NCARDRS, and
- Regulation 32 of the National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013 (the **Regulations**) - to exercise such systems delivery functions of the Secretary of State as are necessary for NHS Digital to deliver the Information System.

The data collected and processed for the Information System will continue to be collected onto the existing NDRS IT systems which will be operated and maintained by UK Health Security Agency (**UKHSA**) from 1 October 2021, on behalf of NHS Digital as a processor for NHS Digital. NHS Digital will be the Controller of the data on receipt on to the NDRS IT systems. It is not envisaged that data will be collected and processed onto NHS Digital infrastructure for 12-24 months following the 1 October 2021 transfer of functions.

Therefore, a Memorandum of Understanding (**MoU**) between NHS Digital and UKHSA supported by a Data Processing Agreement (DPA) with UKHSA as NHS Digital's Processor is in place for the hosting of the data and to enable NDRS staff that have transferred to NHS Digital to continue to access and process the data.

3. National Cancer Registration and Analysis Service

3.1 Background

The National Cancer Registration and Analysis Service (**NCRAS**) is responsible for population-based cancer registration in England. Anyone diagnosed or treated for cancer in England is registered. People suspected of having cancer, people with a high genetic risk of cancer, or people with certain conditions that may lead to cancer are also registered.

Every year the registry collects data on over 300,000 recently diagnosed cancer patients. The registry receives around 25 million records each year from 162 healthcare providers. As well as recent diagnoses, the registry collects data on the continuing treatment of patients diagnosed in previous years. The registry has records of cancer patients going back to the 1940s and the start of the National Health Service (**NHS**).

The NCRAS and the historical data collected will be transferred to NHS Digital on the closure of PHE.

3.2 Scope

Cancer data is processed in order to help the NHS in England, researchers, charities, people with cancer and the public understand what is happening with cancer in this country. To achieve the purposes set out in section 2 of the Direction, the data collected will be used to:

- operate a system dedicated to the surveillance and analysis of the causes, conditions, risk factors, treatments and genetic determinants of cancer;
- count how many cancers are being diagnosed in the population each year (incidence) and how this varies over time or by geographical area;
- calculate how long people are living with cancer compared to in previous years and in other countries (survival) and how this varies over time or by geographical area;
- measure and improve the accuracy of the stage of cancer which people are diagnosed at, and how this varies over time or by geographical area;
- measure and improve the quality of cancer diagnostics, treatments and services in hospitals across England;
- measure and analyse the differences in treatments and outcomes between patients with cancer and other diseases such as cardiovascular disease.
- reduce variation and inequalities in cancer diagnostics, treatments and services across England;
- understand the end-to-end pathway for people diagnosed with cancer;
- understand the links between risk factors and frequency of specific cancers;

- enable identification of cohorts at a higher risk of cancer to whom screening interventions can be targeted;
- understand the links between factors in the tumour and patient, and cancer outcomes;
- monitor and assist with improvements to national audits and cancer screening programmes; and
- provide vital information to NHS Clinical Genetics Services for patient diagnosis, management and counselling, and to improve interpretation of genomic findings in clinical practice.

The collection aims to provide the following benefits to patients and the public:

- to increase prevention and early diagnosis of cancer;
- to improve the management of NHS cancer services;
- to improve NHS cancer treatment and care; and
- to improve patient outcomes, including better quality of life and longer survival.

Once a person is registered, the service will continue to receive information about their cancer whenever they are seen or treated for cancer. Data collection will continue up to, and will include details of, an individual's death or earlier embarkation from England and Wales (collection will recommence if the patient returns). The registry employs an event-based registration model. A cancer registration record will be made at patient and tumour level. In other words, a patient with more than one tumour will have more than one set of tumour records.

A wide range of information will be received and used to develop a rich data resource over time. The data will be accessible to those authorised through the Cancer Analysis System (**CAS**), which is able to create a monthly snapshot of data for analysis. A provisional quarterly snapshot will be generated to facilitate timely data release, and then a designated annual snapshot when each calendar year of registrations is completed. The annual snapshot will be used to produce national statistics.

3.3 Source

The data collection covers individuals in England diagnosed with cancer; treated for cancer; suspected of having cancer; and with certain conditions that may lead to cancer.

NCRAS will process a number of linked and curated datasets to form a cancer registration record. The data will be received from multiple sources both external to NHS Digital and from data already held by NHS Digital.

Following the transfer of functions from Public Health England to NHS Digital on 1 October 2021, external providers of the data will continue to make their submissions in the same way. However, the receiving point will be NDRS IT systems hosted by UKHSA, acting as a Processor on behalf of NHS Digital. NHS Digital will be the Controller of the data on receipt.

Most of the data that will be collected under the Directions comes from NHS healthcare settings, including (but not exclusively) NHS organisations that provide the following services:

- Laboratory databases – including pathology, haemato-oncology diagnostic systems and genomic laboratory information management systems (**LIMS**);
- cancer services and oncology systems;
- hospital patient administration systems;
- diagnostic imaging reports.

Data used to form a cancer registration record that will be collected under the Directions includes the following to achieve the purposes set out at section 2 of the Direction:

1. **Cancer Outcomes and Services Data set (COSD):** Data from providers of NHS care about all patients diagnosed with cancer or receiving cancer treatment in, or funded by the NHS in, England.
2. **Systemic Anti-Cancer Therapy (SACT) Data:** Data from NHS Trusts in England about all anti-cancer drugs given to adult and paediatric patients with solid or haematological malignancies. This will include patients from Wales receiving treatment in England.
3. **Radiotherapy Data Set (RTDS):** Data from providers of radiotherapy services to the NHS about every patient receiving relevant types of radiotherapy exposures. This will include patients from Wales receiving treatment in England.
4. **Somatic Dataset:** Data from NHS Genomic Laboratory Hubs, Molecular Pathology and Haematology laboratories for somatic molecular and genomic tests performed directly on tumour tissues for diagnostic, prognostic, disease/treatment monitoring and personalised medicine purposes.
5. **Germline Dataset**: Data from NHS Genomic Laboratory Hubs for germline molecular and genomic tests performed on individuals in the context of hereditary cancer predisposition syndromes (e.g. BRCA1/2), including full screen diagnostic tests on symptomatic individuals and predictive/presymptomatic cascade tests for relevant family members; and germline tests performed in the context of personalised medicine for cancer patients e.g. DPYD testing to identify individuals at elevated risk of drug toxicity.

3.4 Creating a cancer registration record

The following data items will also be collected under the National Disease Registries Directions 2021 or are currently collected by NHS Digital under other section 254 directions or via other data sharing arrangements. The data items below will be linked to the data in section 3.3 (Source) to create a cancer registration record and include, but are not limited to:

1. **Cancer Waiting Times:** collected by NHS Digital pursuant to the National Cancer Waiting Times Monitoring (SCCI0147) Directions 2016 issued by NHS England. The data is collected in accordance with DCB0147: National Cancer Waiting Times Monitoring Data Set from acute trusts, care trusts and contracted independent sector providers that deliver cancer outpatient, cancer inpatient, cancer screening or cancer treatment services.

2. **Primary Care Prescriptions Data:** data about medicines dispensed or supplied in primary care collected under the Directions from the NHS Business Services Authority (NHSBSA). In practice, until the cancer data is collected and processed on NHS Digital infrastructure the data will be provided to NHS Digital by UKHSA, who also need this data flow from the NHSBSA. This flow will be in accordance with an agreement between NHSBSA and UKHSA which authorises UKHSA to share the data with NHS Digital.
3. **Hospital Episode Statistics:** Data derived from data submitted by providers of NHS services to the Secondary Use Service (**SUS**), which is delivered by NHS Digital pursuant to the Spine Services (No 2) 2014 Directions issued by the Secretary of State.
4. **SUS Payment by Results**: collected by NHS Digital under the Spine Services (No 2) 2014 Directions from providers of NHS services. The extract used for NCRAS is based on admitted patient care episodes for cancer related registerable conditions only. SUS data provides details of treatment, particularly surgery. The data is provided by NHS Digital to UKHSA and therefore initially will be accessible to NDRS on the UKHSA infrastructure.
5. **Diagnostic Imaging Data:** collected by NHS Digital pursuant to the Establishment of Information Systems for NHS Services: Diagnostic Imaging Data Set Directions 2016 issued by NHS England from providers of diagnostic imaging tests carried out on NHS patients.
6. **Death Registration Data:** mortality data collected under the Directions from the Office for National Statistics (**ONS**). In practice, until the cancer data is collected and processed on NHS Digital infrastructure the data will be provided to NHS Digital by UKHSA, who also need this data flow from the ONS. This flow will be in accordance with an agreement between ONS and UKHSA which authorises UKHSA to share the data with NHS Digital.

As cancer registration and research continues to develop, it is important that NCRAS has the permissions to explore and access new data opportunities where they arise. These may include primary care data, historic consented studies where consent materials do not preclude data sharing with a third party, existing patient / clinical registers where there are explicit consent processes in place, and other data sources that NCRAS may be currently unaware of, do not yet exist or NCRAS does not have access to, where sharing of data would be deemed appropriate within the interpretation of the range of permissions laid out in this document. In the event that there is a proposal to add new data sources to the NCRAS, the Change Control process at section 9 will be followed to ensure any data collection proposals are aligned with the purposes set out at section 2 of the National Disease Registries Directions 2021.

3.5 Category

A cancer registration record contains data items such as:

- **Personal data**, such as name, gender, sex, date of birth, age, address and postcode, height/weight, BMI, NHS number, local identifier (e.g. hospital number or specimen number), and habits relating to smoking and alcohol consumption;

- **Special categories of personal data**, such as ethnic origin, sexual orientation, health data (diagnosis and treatment of cancer) and genetic data.

For each tumour, the dataset aims to record a cancer patient's:

- referral
- diagnosis
- the type of cancer they have
- the characteristics of their tumour
    - including biomarker status and somatic test results
- specific clinical and pathological information about certain diseases
- the hospitals and centres where they are seen and treated
- the healthcare professionals in charge of their care
- the treatment and care they receive
- and their outcomes

The primary patient identifier is the NHS number, but date of birth, full name and address are also used for patient identification and linkage.

3.6 Frequency

1. **Cancer Outcomes and Services Data Set:** Data is submitted in accordance with DCB1521: Cancer Outcomes and Services Data Set. This is a monthly data collection in line with the ISN. Reported on or around the 25-working day post the end of each month.
2. **Systemic Anti-Cancer Therapy Data:** Data is submitted in accordance with DCB1533: Systemic Anti-Cancer Therapy Data Set. This is a monthly data collection to be reported by the last day of each month.
3. **Radiotherapy Data Set:** Data is submitted in accordance with SCCI0111: Radiotherapy Data Set and http://digital.nhs.uk/isce/publication/dapb0111. This is a monthly data collection reported on the twentieth working day of each month.
4. **Somatic Dataset:** Data is submitted monthly from NHS Genomic Laboratory Hubs, Molecular Pathology and Haematology laboratories
5. **Germline Dataset:** Data is submitted monthly from NHS Genomic Laboratory Hubs

4. Analysis

4.1 Data collected for cancer registration and analysis purposes

A cancer registration record is made at the patient and tumour level. Data submitted will be reviewed by skilled cancer registration officers (**CROs**) with the assistance of some automated tools for data linkage and de-duplication of identical data sources. Review will include manual extraction of data from text-based pathology reports. CROs require detailed knowledge of cancer biology, coding and terminology. As cancer care is often delivered at different hospitals, information may relate to patient episodes at many organisations, which can introduce inconsistencies in the incoming data. The CRO will examine all data sources

and, if necessary, seek additional source information from primary or secondary care via correspondence or, where agreed with the relevant provider of the data (the **Controller**), direct interrogation of hospital radiology or electronic health records via remote secure access in accordance with a data processing agreement (NHSD as Processor) with the relevant Controller.

Under the Directions and this Requirements Specification, the information to be sourced from the datasets listed in Source (section 3.3 above) must be provided in accordance with the Cancer Outcomes and Services Dataset (COSD), the national standard for reporting cancer in England, or any subsequently approved version of the same standard. During the development of a revised COSD standard, this Requirements Specification will also be reviewed to ensure any updates to the standard are aligned with the National Disease Registries Directions 2021 (see also the Change Control process at section 9 below).

4.2 Data collected for analysis purposes only

In addition to the linked and curated datasets that make up a cancer registration record described in section 3.3 and section 3.4, **the following datasets will also be collected under the National Disease Registries Directions 2021** for analysis purposes:

- **Cancer Drugs Fund (CDF) data**: NHS England and NHS Improvement's Blueteq database captures the CDF population. The Blueteq data is collected under the Directions from NHS England and NHS Improvement for CDF evaluation purposes.
- **Quality of Life (QoL) data**: The Cancer QoL Survey is commissioned by NHS England and NHS Improvement and the data is collected by NHS Digital under the Directions. The survey is for people in England who have been diagnosed with cancer, aiming to measure how quality of life may have been affected for people diagnosed with cancer. A survey provider (currently Quality Health) is formally engaged as a Processor to NHS Digital, and manages the survey invite and response system. NHS Digital will provide the survey provider with a regular extract of eligible participants. Those who agree to participate provide their consent for their information to be collected and stored in the CAS for the reporting of survey results and linkage with information related to each person's diagnosis and treatment.
- **National Cancer Patient Experience Survey (NCPES)**: the NCPES is commissioned and managed by NHS England and NHS Improvement. The implementation, analyses and reporting is completed by Picker. The data is collected by NHS Digital under the Directions. Patients are informed that their information will be shared with NDRS at NHS Digital, see: "What happens to my answers".
- **National Cancer Diagnosis Audit (NCDA) data**: The NCDA processes data collected under the Directions about patients diagnosed with a new primary cancer. The audit is conducted in cycles: patients are identified through NCRAS based on their GP practice at the time of diagnosis. The diagnosis date and type of cancer is made available to verified GPs who can log in to a secure online portal (the Audit portal) and add data including the nature of the referral, use of investigations ordered by primary care, and symptoms which the patient presented with. This allows NCRAS

to analyse the data to audit the referral of cancer patients and the interaction of primary and secondary care to promote earlier and swifter diagnosis.

- **Post Colonoscopy Colorectal Cancer (PCCRC) audit data**: The PCCRC is an audit funded by Bowel Cancer UK and run by NCRAS in partnership with Health Data Insight. It uses the Audit portal developed for the NCDA. Data is collected under the Directions from NHS Trusts who enter additional information on colonoscopies which were negative for cancer, but where the patient was later diagnosed with PCCRC. The purpose of the audit is to identify improvements that can be made in the colonoscopy process to reduce the rate of PCCRCs. In the audit portal, participating clinical leads access a list of their patients, identified by NCRAS through linkage of HES, bowel cancer screening, and cancer registry datasets. Selecting one of these cases will lead to the audit template.
- **Rapid Diagnostic Centre (RDC) dataset**: the RDC minimum dataset is collected under the Directions by NCRAS from NHS Trusts via the secure EnCORE API portal (see section 8.1 below) on behalf of NHS England and NHS Improvement (NHSE&I). The data is used by NHS Digital in the Trusted Research Environment (TRE), with no further involvement by NCRAS. The RDC minimum dataset has been defined by NHSE&I and can be found here: https://www.england.nhs.uk/wp-content/uploads/2019/07/rdc-vision-and-1920-implementation-specification.pdf
- **Retinoblastoma Register and Audit**: data about children with retinoblastoma (RB) is collected under the Directions from the two RB treatment centres in the UK (Birmingham Women's and Children's NHS Foundation Trust and Barts Health NHS Trust). NCRAS provides a bespoke national disease register and audit for the collection, audit and analysis of data which will replace the existing in-house databases used by the RB treatment centres. The register will enhance, expand and unify the current data collection from the RB clinical teams by combining it with data available from cancer registration sources (COSD, SACT, MDT, RT, Pathology). It will provide a clinical research / audit resource to evaluate the quality of the services being provided, the effectiveness of treatment strategies and the short term and longer-term health and wellbeing of patients following completion of therapy, in relation to treatment-related toxicity.
- **National Breast Screening System**: data collected under the Directions from every breast screening service in England. Data is extracted through a secure, encrypted connection into the Screening Histories Information Manager (SHIM), which is an NHS Digital system hosted by UKHSA. The data is linked with the cancer registration data and the purpose is to simplify and improve the identification of interval cancers in the NHS Breast Screening Programme and automate most of the work required to find out the screening status for each breast tumour i.e. was the tumour detected through breast screening.
- **SARS-COV-2 subset of Second Generation Surveillance System (SGSS) dataset:** data collected under the Directions from the UKHSA. This is demographic and diagnostic information from laboratory test reports for patients tested for the suspected and confirmed causative agent for COVID-19, as recorded in the SGSS. The SARS-COV-2 data is submitted weekly by UKHSA and is composed of test

identifier (e.g. barcode) and patient identity, laboratory and requestor details, patient testing details and results, and deduplication identifiers.

- **UK Transplant Registry (UKTR) organ transplant data:** data collected under Directions under the Directions from NHS Blood and Transplant (NHSBT). Organ transplant recipients (OTRs)_will be matched against the NCRAS by NHSBT sending pseudonymised identifiers to NHSD. Once received, NCRAS will then analyse the data to identify OTRs with cancer. This is with a view to better understanding the relationship between the immune suppression which occurs as a result of being an OTR and skin cancer and to improve future research. Further, subpopulations of at-risk patients can be targeted for skin cancer prevention programmes and closer surveillance.
- **National Institute of Cardiovascular Outcomes Research (NICOR) data:** cardiovascular data will be linked to cancer registration data to enable the Virtual Cardio-Oncology Research Initiative (VICORI) programme to investigate differences in treatments and outcomes between patients with cancer, cardiovascular disease, or both. The programme has a series of aims including to investigate whether patients with a cancer diagnosis, who subsequently develop cardiovascular disease, are managed differently from patients without cancer, whether cardiovascular treatments, interventions, and surgery alter a patient's subsequent risk of developing cancer, whether cancer treatments affect the risk of long term adverse cardiovascular disease states or events, and whether patients with cardiovascular disease, who subsequently develop cancer, are managed differently from patients without cardiovascular disease.

4.3 Quality assurance

The registration process will have embedded robust automated and manual quality control and assurance checks, both at the individual record level and for the set of registrations. As records are added to the Information System, input validation checks will compare them with expected values and cross-validate against other records, for example if treatment occurred after death. Quality control checks will be undertaken during processing and inconsistencies resolved. Checks across finalised records will be performed before the quarterly sign-off of data for release. Trend and population-level checks will also be performed.

Once the data has gone through system validation checks it will be allocated to registration officers for further processing. Registration officers will use the data to:

- create new cancer registrations;
- amend and update existing ones; and
- check that registration records are complete and high quality.

If necessary, NCRAS staff will review primary sources of the data to find vital and missing information. They may do this by contacting staff in NHS cancer services, visiting hospitals, or, where agreed with the relevant Trust, by remotely logging on to hospital systems in accordance with a data processing agreement (NHSD as Processor) with the relevant Controller.

Cancer registry data quality will be assessed in terms of completeness, validity, comparability and timeliness of submissions. There are extensive published quantitative data which demonstrate quality, for example the performance indicators published by the United Kingdom and Ireland Association of Cancer Registries (**UKIACR**) and the Cancer Incidence in Five Continents (**CI5**) publications. English cancer registration data has previously been assessed for completeness and accuracy against randomized controlled trials data, inpatient hospital data and simulations, with positive results. Validity, comparability and timeliness of the registry data will continue to be tested and published alongside the national statistics of both cancer incidence and survival.

4.4 Derivation of key diagnostic information

Derivation of key diagnostic information: incidence date and stage at diagnosis

Date of incidence will be defined according to the European Network of Cancer Registries (**ENCR**) rules. Key events from which the registry will derive the date, include: pathological verification; discussion at a Multi-Disciplinary Team (**MDT**); referral with suspicion of cancer; and diagnostic imaging.

Staging information received will include: data from MDTs; pathology reports of biopsies and surgical treatment; imaging results; and post-mortems. This information will be reviewed by CROs and collated into a consistent registry incidence date and a single registry-defined stage at diagnosis.

The registry-defined stage will combine all relevant information available to give a single anatomical stage at diagnosis, using the TNM (Tumour Node Metastases) classification system where this is available for a tumour, or a site-specific staging system. In some circumstances the registry-defined stage differs from other stages recorded by a hospital. For example, where patients are managed at multiple hospitals, different components of the staging information may be received from each hospital.

4.5 Data outputs

The cancer registry data will be processed via a custom-built application called EnCORE (English National Cancer Online Registration Environment) and stored in an Oracle database. **These NDRS applications and the data are hosted by UKHSA. UKHSA will provide access under an MoU and Data Processing Agreement with NHS Digital to authorised staff until such time as the applications and data are transferred to NHS Digital infrastructure**. EnCORE is a live application. Monthly snapshots of the data will be cloned into the CAS for analysis and data release. Full details about the data items available in the CAS can be found in the published data dictionary which is updated periodically.

The NCRAS dataset will comprise three main tables: the tumour table holds information about each primary tumour diagnosed in an individual; the patient table holds information about the individuals who are diagnosed with cancer; the treatment table holds the treatment received by the patient. Several data items in the tables are derived using algorithms and multiple data sources, for example the Charlson Comorbidity Index score derived using hospital admission data.

4.6 Other linkage and analysis

NCRAS data may continue to be used nationally and internationally, where there is a lawful basis to do so (and subject to a process of consultation, necessity and proportionality), for a wide range of functions including public health, commissioning, evaluating clinical performance, health care services and research, provided that the purposes are compatible with those set out in the Directions. All new processing activities to be considered in this vein will be subject to an appropriate internal approval process.

5. Consultation

In accordance with section 258 of the 2012 Act, before establishing the Information System, NHS Digital has consulted with:

- The Department of Health and Social Care (**DHSC**)
- Public Health England
- NHSX
- NHS England and NHS Improvement
- The UK Health Security Agency
- Representatives of those that will use the data - current recipients of the data have been informed of the change in Controller and been advised that existing data sharing agreements have been novated to NHS Digital
- Representatives of those from the information will be collected - providers of the data are being informed through existing NDRS communications routes that NHS Digital will now collect the data, this will be supported by communications advising providers of the Data Provision Notice issued by NHS Digital under section 259 of the 2012 Act.

Information Standards

The collection of key NCRAS datasets such as COSD, SACT and RTDS are subject to consultation with providers as part of the Data Alliance Partnership Board information standard process. Consultation over each new version of these datasets involves clinical teams, cancer services, software suppliers, cancer vanguards, commissioners and other stakeholders. During the consultation process, this Requirements Specification will be reviewed and updated where necessary in accordance with the Change Control process in section 9 below.

6. Dissemination/Sharing

6.1 Regular Dissemination/Sharing

6.1.1 De-identified data: CancerStats2 reporting platform

NCRAS will continue to provide rapid operational feedback to hospitals (medical professionals and cancer MDTs) and commissioners about the quality and completeness of the data that have been submitted, to improve the data and outcomes for cancer patients. Interactive online reports will be provided on the CancerStats2 reporting platform which is available on the secure N3/HSCN network to authorised users within the NHS.

CancerStats2 is a repository of information from datasets managed or supported by NCRAS. It is a secure platform which enables users to generate reports from anonymous aggregate NCRAS data on a self-service basis. Reporting tools use **aggregate and anonymised** data.

The process for granting access to the CancerStats2 platform is overseen by the NDRS Associate Caldicott Guardian. When a user requests access to the platform they are required to complete a registration form and agree to a 'User Declaration' regarding the user's data protection and data security responsibilities.

6.1.2 Data shared for National Cancer Audits

NCRAS is the data collection partner for three national audits commissioned by the Healthcare Quality Improvement Partnership (**HQIP**) as part of the National Clinical Audit Patient Outcomes Programme (**NCAPOP**). NCRAS provides English data under data sharing contracts from the following datasets to the respective National Audit teams (which include but are not limited to those set out in this section 6.1.2 and below at section 6.1.3):

- Cancer Registrations
- Cancer Outcome and Services Dataset (COSD)
- Rapid Cancer Registration Dataset (RCRD)
- Hospital Episode Statistics (HES)
- Radiotherapy Dataset (RTDS)
- Systemic Anti-Cancer Therapy (SACT)
- Office for National Statistics (ONS) mortality data
- SARS-COV-2 subset of Second Generation Surveillance System (SGSS) dataset

Each National Audit is contracted by HQIP to produce an Annual Report, and other outputs are published throughout the year as summarised below. The audits also publish Level 2, Level 3 and quarterly reports on CancerStats2 (see 6.1.1).

National Lung Cancer Audit (NLCA): The NLCA (previously named LUCADA) is run by the Royal College of Physicians and commenced in February 2015. The cohort includes men and women of all ages diagnosed with lung cancer in England, Wales (and more recently Jersey and Guernsey).

Project Status Reports created by NHS Digital will be shared with Audit stakeholders, including the Royal College of Physicians, on a quarterly basis. Prior to the contracted annual report a data snapshot is taken for trust validation.

National Prostate Cancer Audit (NPCA): The NPCA is run by the Royal College of Surgeons and commenced in April 2013. The cohort includes men of all ages diagnosed with prostate cancer in England and Wales.

Project Status Reports created by NHS Digital will be shared with Audit stakeholders, including the Royal College of Surgeons, on a quarterly basis. Prior to the contracted annual report a data snapshot is taken for trust validation.

National Audit of Breast Cancer in Older Patients (NABCOP): The NABCOP is run by the Royal College of Surgeons and commenced in April 2016. The cohort includes women aged 50 years and above diagnosed with breast cancer in England and Wales.

In addition to the above, the NABCOP has received data from NCRAS in relation to the National Cancer Patient Experience Survey (NCPES) (see section 4.2 above); Primary Care Prescriptions Data and the NHS Breast Screening Programme/Association of Breast Surgery (**NHSBSP/ABS**) Screening Audit dataset.

Project Status Reports created by NHS Digital will be shared with Audit stakeholders, including the Royal College of Surgeons, on a quarterly basis. Prior to the contracted annual report a data snapshot is taken for trust validation.

6.1.3 Other regular data sharing/dissemination

National Cancer Diagnosis Audit (NCDA): patients are identified through NCRAS based on their GP practice at the time of diagnosis. The diagnosis date and type of cancer is made available to verified GPs who can log in to the secure online Audit portal and add data including the nature of the referral, use of investigations ordered by primary care, and symptoms with which the patient presented.

Post Colonoscopy Colorectal Cancer (PCCRC) audit: patients are identified through NCRAS based on patients diagnosed with colorectal cancer who previous had colonoscopies which were negative for cancer. A list of NHS numbers is made available in the secure Audit portal to the named clinical lead in trusts where the colonoscopies identified by NCRAS were undertaken. In the Audit portal, clinical leads access a list of their patients. Selecting one of these cases will lead to the audit template. As the audit progresses, aggregated data will be shared with all trusts. There will be no sharing of patient or endoscopist level data other than to the trust who carried out the endoscopy.

Retinoblastoma Register and Audit: clinicians specialising in this rare childhood cancer can securely access the data NCRAS has collected about their patients.

Cancer Drugs Fund (CDF) reports: NCRAS has a partnership with NHSE on cancer data, this includes providing an evaluation of drugs that have entered the CDF, to support decision making by the National Institute for Health and Care Excellence (NICE) committees. NCRAS shares aggregate information about CDF indications with NHSE/I, NICE and the relevant pharmaceutical company in the form of quarterly, annual and final (end of CDF) reports. Final reports are published by NICE but quarterly and annual reports are never published. The data sharing/dissemination is governed by an overarching Privacy Impact Assessment with a modified NICE committee form.

Breast Screening after Radiotherapy (BARD): BARD is the first national targeted screening programme in England and aims to save the lives of women who are at risk of breast cancer following radiotherapy treatment to breast tissue when under age 36 years. This personalised care programme achieves this aim through detailed analysis of radiotherapy treatment data, identification of patients at risk of breast cancer and eligible for enhanced annual breast screening, and subsequent completion and submission of referrals for these patients. This enhanced breast screening aims to identify breast cancers in their

early stages, when they are much more treatable and survival outcomes are more favourable. NCRAS data is used to identify cohorts of women at high risk. To enable this purpose, data are shared with the NHS Breast Screening programme, patients' GPs and NHS radiotherapy providers.

Cancer Services Profiles: These are currently published on the PHE Fingertips Tool which is being transferred to UKHSA. Therefore, the following information will be disseminated to UKHSA to enable the publication - 31 indicators on demographics; cancer screening; Two Week Wait referrals; diagnostic services; emergency presentations and admissions by GP practices, CCG and at National level (Official Statistic).

Screening programmes: Data is shared with NHSE/I to support implementation, auditing and to measure the impact of the national cancer screening programmes. This includes, but is not limited to, data sharing to support quality assurance and service evaluation of interval cancer and invasive cancer analysis in breast, bowel and cervical screening, including the Screening Histories Information Management System (SHIM) for interval breast cancers.

SARS-COV-2 subset of SGSS: NHS Digital will only disseminate SARS-COV-2 SGSS data collected from UKHSA where the information is linked to other information controlled by NHS Digital, with the exception of SARS-COV-2 information collected from SGSS for residents of the Devolved Administrations which may be disseminated directly by NHS Digital to the relevant Devolved Administration Public Health Agency.

6.1.4 Ad-hoc sharing of anonymous/aggregate data

NCRAS will continue to make cancer data available for health research purposes so that improvements in treatment and outcomes can be realised faster. Information about how to access the data can be found at https://digital.nhs.uk/ndrs/data/access-to-data.

NCRAS provides a central point for receiving, managing and tracking ad-hoc analytical queries. Many of the enquiries received by NCRAS can be answered using the wealth of information that has already been published or released. If some specific analysis (which is not readily available) needs to be produced, it will be assessed in terms of whether the analysis can legally be released, and whether there is capacity in the team to undertake the work.

NCRAS also contribute to a wide range of international cancer studies, including but not limited to, the International Cancer Benchmarking Partnership with the International Agency for Research on Cancer / World Health Organisation. Taking part in these studies greatly raise the profile of NCRAS, its data and expertise, as well as often allowing for opportunities to improve information based on cross-examination of our data with other registries.

All data releases are sent with a completed cover sheet to provide an overview of what, when and who produced the analysis; this may also include the SQL code used to extract the cohort and any other code used to analyse the data.

NCRAS also deals with Parliamentary Questions and Freedom of Information requests relating to cancer data and analysis and prioritises resource to respond to these as appropriate.

6.2 Applications for data

NHS Digital will use its statutory powers, including in regulations 2 and 3 of COPI and section 261 of the 2012 Act to disseminate information to applicants that have been approved to receive or access the data in the form requested. All applications for data will be assessed to determine whether the applicant's needs can be met with anonymous data or whether the project protocol can be revised in a way that allows for anonymous data to be used.

6.2.1 Data Access Request Service (DARS)

All applications will be managed via the Data Access Request Service (**DARS**), which will include oversight by the Advisory Group for Data where appropriate.

More information about the data sets and collections that NHS Digital hold and that may be disseminated or used for linkage where approved can be found on the NHS Digital Data Collections and Data Sets webpage.

Any dissemination of the data through DARS will be published in the NHS Digital Data Release Register.

6.2.2 Requests from NHS clinical services for confirmation of cancer status (Genetics requests)

NCRAS will continue to support NHS Clinical Genetics, and other clinical services, by offering a family history diagnosis checking service. Individuals with a strong family history of particular types of cancer, often at an unusually young age, attend cancer genetic counselling clinics in order that a genetic cause for their cancer predisposition can be investigated, and appropriate testing or screening can be offered.

All requests for England and Wales are managed through the online Genetic Information Request Portal. NHS Digital deals with requests from all English centres and Public Health Wales deals with requests from Welsh centres. There are data processing agreements in place that allow NHS Digital to provide returns on Welsh residents and vice versa. There are a handful of requests that cannot go through the portal as registering for access requires an nhs.net email address, this would include overseas requests for approved services in Australia and New Zealand – these are dealt with using secure email transfer/Egress and details of all information exchanged via this route are recorded to provide an audit trail.

Clinical Geneticists and Genetic Counsellors working in registered NHS Clinical Genetics or family cancer services can access a secure online portal on the NHS network, into which they can input details of their patient's relative(s), along with signed consent in the case of living family members. Access to the portal is only given to approved users within approved/known Clinical Genetics or Family History services – this list is reviewed annually.

The requestors complete a number of key fields to allow a match to be found in EnCORE – these include patient demographics, consultant, family/proband identifier from service etc.

For patients who are alive, a consent form or power of attorney must be submitted; these are uploaded to the portal and the requestor must click a checkbox to confirm that they have

consent to request this information. Consent forms are not checked for every single case but as part of due diligence, NHS Digital will continue the process of carrying out an audit to check 5-10% of all consent forms across all submitting Trusts. If any issues or non-compliances are identified, then a deeper audit of the Trust in question will be undertaken to determine whether this was a one-off error (i.e. human error) or a more systematic issue. All non-compliances will be fed back to Trusts and they will be informed to notify their IG teams.

NCRAS matches each person's details to their record in the cancer registration dataset and returns information on the correct diagnosis to the clinical genetics service, via a secure NHS email account. The actual requests are then dealt with by the Genetic Processing Officers who verify the details are a correct match before a response is returned. The specific data variables that are returned include:

- Requestor details
- Patient demographics
- Diagnosis
- Date of diagnosis
- Hospital where treated
- Tumour details:
    - Status (final, provisional etc.)
    - Hospitals attended
    - Postcode at diagnosis
    - Diagnosis date
    - Site
    - Histology
    - Laterality
    - Molecular (included where available)

This service enables the clinical geneticist or genetic counsellor to have complete confidence of the diagnoses within the family, and thus to be able to give personalised genetic counselling and to offer the most appropriate management to their patient. Confirmation of family history is a pre-requisite before genomic testing can be offered to the individual. Around 20,000 requests are received each year, with an average response rate of 3 working days.

6.3 National data opt-out

National data opt-outs will not apply to the collection of the data for the cancer registry because to fulfil the National Disease Registries Directions 2021, NHS Digital will issue a Data Provision Notice (**DPN**) to the holders of the data, in accordance with section 259 of the 2012 Act. A DPN imposes a legal obligation on the provider to supply the data to NHS Digital and the national data opt-out policy does not apply where there is a legal obligation to share data.

National data opt-outs will be applied in accordance with the national data opt-out policy to the dissemination of data from the NCRAS via DARS to restrict the dissemination of an

individual's confidential patient information by NHS Digital where the national data opt-out applies.

More information about the national data opt-out is available at:
https://digital.nhs.uk/services/national-data-opt-out

6.4 Cancer Registry opt-out

Patients are able to opt-out of their personal data being collected into NCRAS as a matter of DHSC policy. NCRAS has a Standard Operating Procedure that describes the process that must be followed, including how a request to opt-out should be handled, the records that are kept regarding the request, and the technical measures in place to ensure subject data is removed. The NDRS Associate Caldicott Guardian is responsible for engaging directly with the individual to discuss the value of cancer registration to healthcare services and the possible consequences of opting out.

Where an individual has opted out from NCRAS, information provided by the holder of the data will not be processed and will be deleted. In accordance with this opt-out, a request by a patient for their details to be removed from the registry will also be upheld. A secure opt-out list is maintained to ensure that those who have opted out are not inadvertently registered or re-registered.

NHS Digital is directed to put measures in place as part of the establishment and operation of the Information System to ensure that where an individual opts out of their personal data being held or a person with parental responsibility for a child opts out of that child's personal data being held no new data about them is collected.

More information about the Cancer Registry opt-out is available at:
https://www.ndrs.nhs.uk/national-disease-registration-service/patients/opting-out/

7. Publication

7.1 Data to be published

As a core part of its remit, NCRAS will continue to produce a variety of reports and publications using cancer registration data, including national statistics, peer-reviewed articles in scientific journals, data reports, data tables, and cancer information tools.

A Standard Operating Procedure describes the planning, approval, sign-off and publication of documents and outputs produced by NCRAS. It includes guidance on how the use of NCRAS data should be formally acknowledged in publications.

Regular publications and formal Open Data (public) releases, and releases as a result of enquiries (see 6.1.4 Ad-hoc sharing of anonymous/aggregate data) will be managed in accordance with the NCRAS Data Release Guide which describes the assessment required to be undertaken for each data release where a contract or data sharing agreement is not being used to cover the release.

For every row-level data release without a contract, or where there could be any doubt about whether an aggregate data release is anonymous, a Data Protection Impact Assessment will

be undertaken, which will then be considered for sign off by the NDRS Associate Caldicott Guardian.

7.1.1 Regular publications

A range of NCRAS outputs will continue to be published and updated on a regular basis at http://www.ncin.org.uk/publications/:

Annual

- Cancer incidence: The numbers and rates of newly diagnosed malignant primary neoplasms (cancer tumours) registered each year (National Statistic)
- Cancer mortality: The numbers and rates of people dying from cancer each year (National Statistic)
- Cancer prevalence: The number of people who are alive on a specified date and have previously been diagnosed with cancer
- Cancer Survival Statistics: England, Adults and Childhood; Geographic Patterns of cancer survival; survival by stage; survival index for clinical commissioning groups (National Statistic)
- Routes to Diagnosis: Cancer incidence breakdowns categorised by the different diagnostic routes to a cancer diagnosis
- Treatment: Population-based statistics on the patients recorded to have received chemotherapy, radiotherapy and surgical tumour resections for their tumour
- Stage by CCG: Detailed stage breakdown by cancer site and CCG for all cancers
- 30-day mortality after Systemic Anti-Cancer Therapy: 30-day post chemotherapy mortality for patients with certain cancers treated with curative or palliative intent
- Childhood cancer statistics: Incidence, survival, prevalence and mortality for cancer diagnosed among children under the age of 15 resident in England/UK
- Detailed Statistics: Statistics on incidence, survival, treatment and routes to diagnosis for small groups of cancer patients produced by the Get Data Out team https://www.ndrs.nhs.uk/get-data-out/
- Median Pathway: Median pathway analysis by patient demographics, stage at diagnosis, route to diagnosis, and geography
- Urgent suspected cancer referrals conversion and detection rates: Cancer Waiting Times (CWT) urgent suspected cancer referrals: referral, conversion and detection rates

Quarterly

- 75% stage monitoring: Measuring progress against the Long-Term Plan ambition to diagnose 75% of all cancers at an early stage
- Emergency Presentations: Estimated proportion of all malignant cancers where patients first presented as an emergency (Official Statistic)

Monthly

- Radiotherapy activity: Radiotherapy activity (episodes and attendances) in hospitals for different cancers and the proportion of treatments using intensity modulated radiotherapy;
- Chemotherapy activity: Chemotherapy activity in hospitals for different cancers;
- Rapid Cancer Registration Data: proxy tumour registrations and some associated events on the cancer patient pathway (e.g. surgery, radiotherapy and chemotherapy) from January 2018 to the most recently available data on cancer diagnoses;
- Cancer Alliance reporting: Summary indicator grids containing a range of cancer metrics by Cancer Alliance, Sustainability and Transformation Partnership (STP) and Clinical Commissioning Group (CCG), compared with the national benchmark, expected values or operational standards;
- COVID-19 dashboards: projects to provide near real-time cancer data for England to support clinical and policy decision making in relation to the COVID-19 pandemic.

CancerData

NCRAS will continue to routinely publish anonymous aggregate data publicly at CancerData. The portal includes highly interactive tools such as dashboards, statistics on Covid -19 and radiotherapy, incidence, survival and mortality, treatment, routes to diagnosis and stage at diagnosis. Many of the reports can be broken down Clinical Commissioning Groups, Sustainability and Transformation Partnerships and Cancer Alliances across England, or broken down by population demographics and for different tumour types.

7.1.2 Ad-hoc and one-off publications

NCRAS will continue to publish ad-hoc outputs as part of its core remit, this will include one-off reports and peer-reviewed journal articles. A Project Panel made up of Senior NDRS personnel is responsible for reviewing and approving analytical work, including confirming the sign-off and publication process that should be followed; all of which must be documented in an approved Project Proposal. The NCRAS Standard Operating Procedure and Data Release Guide are followed for every publication.

7.2 Data prohibited from being published

NCRAS has a longstanding partnership with NHS England on cancer data, that includes providing an evaluation of some drugs that have entered the new Cancer Drugs Fund (**CDF**) (July 2016) to support decision making by the National Institute for Health and Care Excellence (**NICE**) committees. There are temporary commercial restrictions on releasing some data relating to some CDF indications. These restrictions apply until NICE make their decision public as to whether a given drug will go into routine commissioning or be de-commissioned; at which point, NCRAS' full evaluation report (redacted by NCRAS if advised by the NDRS Associate Caldicott Guardian to preserve patient confidentiality) will be published on the NICE webpage. Exceptions to this embargo are requests to support direct patient care, if a patient has explicitly consented to their data being shared, or to the manufacturer whose drug is being evaluated.

In accordance with a request from UKHSA, NHS Digital will not publish information collected from SGSS or based on analysis of SGSS data alone, without first obtaining consent to do so from UKHSA. NHS Digital may publish information and analysis derived from linking the SGSS information to other sources of information held by NHS Digital.

8. System Delivery Function

8.1 Collection and storage of the data whilst on UKHSA infrastructure

For a period, data will be submitted to and held on existing NDRS IT systems which will be operated and maintained by UKHSA from 1 October 2021, whereby UKHSA will host the NDRS IT systems and data as a Processor on behalf of NHS Digital as Controller. A Memorandum of Understanding and a Data Processing Agreement are in place to set out the hosting and data access arrangements.

Health and care providers submit data via either via NHSmail to NHSmail or by TLS-protected upload to https://nww.api.encore.nhs.uk (private HSCN network users only). These provide secure, encrypted pathways for the sharing of patient data. Information is also received via secure file transfer from other providers.

The NDRS IT systems that will be hosted by UKHSA are:

The cancer registration system (EnCORE) is backed by an Oracle 11.2 database accessed via a web-based front end written in Ruby on Rails. It is configured across six physical servers. The database is hosted on servers running a Linux operating system.

The Cancer Analysis Server (CAS) is a collection of Oracle databases. Each snapshot begins its lifecycle as a point-in-time snapshot of the live EnCORE registration database. Analysts and Database Administrators (**DBAs**) work together to write an “analysis schema”: a series of SQL views and functions which assist interpretation and analysis of the data. Access to the Analysis server is via a gateway server and requires two-factor authentication. The work that a user can actually do within the system is controlled by database privileges set on the Oracle databases.

Genetic Information Request Portal is an online family history diagnosis checking service accessible by authorised Clinical Geneticists and Genetic Counsellors working in registered NHS Clinical Genetics or family cancer services.

CancerStats2 platform is an online platform used by medical professionals, hospital cancer teams and commissioners to generate reports using NCRAS data on a self-service.

Audit portal for National Cancer Diagnosis Audit (NCDA and Post Colonoscopy Colorectal Cancer audit and the Retinoblastoma register audit (see section 4.2 Data collected for analysis purposes only).

Screening Histories Information Manager (SHIM), which is a data linkage and analysis system, that links data from EnCORE with data collected under the Directions from the National Breast Screening System which is used by every breast screening service in England.

8.2 Collection and storage of the data once on NHS Digital infrastructure

Over time, the data will be stored on NHS Digital's DPS (Data Processing Services) platform. New IT systems and applications may need to be procured in addition to replace some of the systems and applications listed above (unknown at time of writing).

9. Change control process

Changes to this Requirements Specification will be managed and agreed by NHS Digital and an authorised officer of the Department of Health and Social Care on behalf of the Secretary of State, to ensure that such changes remain aligned with the National Disease Registries Directions 2021.