



## PHE cancer data sets, linkage and availability (v1.1)

Public Health England's National Cancer Registration and Analysis Service (NCRAS) is a population-based registry of all cases of cancer diagnosed or treated in England. NCRAS routinely collects both patient and tumour level information to build a rich picture of the cancer patient pathway. This data can also be linked to other data, such as Hospital Episode Statistics to build an incredibly valuable data resource to support public health, healthcare and research. To maximise the utility of this wealth of data PHE has a dedicated process and structure to ensure we can share what we hold with experts in a transparent, fair and lawful manner.

This document describes the data held by NCRAS and contains descriptive information about their linkage attributes and temporality.

**Table 1** describes the data held by NCRAS and for which PHE is the data controller.

**Table 2** describes the data held by NCRAS but where PHE is not the data controller. These data are only available when linked to the data in Table 1 or where PHE has specific agreements in place to support access to these data as discrete datasets.

### Access to PHE data

For more information or advice about cancer data or how to request access to PHE data sets, please visit the [Office for Data Release \(ODR\) website](#) or contact the ODR by email at [ODR@phe.gov.uk](mailto:ODR@phe.gov.uk).

In a continuous effort to improve our guidance, we look forward to your thoughts, questions and feedback regarding the content or format of this document. Please contact the ODR by [email](#).

Table 1: Data sets held by Public Health England:

| Data set  | Description  | Years of data available while linked to the cancer registration tables                      | Years available as a discrete dataset | Minimum level of linkage type |
|---|--|---|---------------------------------------|-------------------------------|
| <b>Cancer registry tables</b>   |  |   |                                       |                               |
| Cancer registration (patient table)   | To achieve comprehensive registration for all registerable tumours, the National Cancer Registration and Analysis Service (NCRAS) brings together data from more than 500 local and regional data sets to build a picture of an individual's treatment from diagnosis. The cancer registration tables include data on the patient, their diagnosis, tumour characteristics and details of the care and treatment received. Most of this data is submitted electronically by NHS providers (such as NHS Trusts) to NCRAS on a monthly basis in line with the requirements of the national standard for reporting cancer data, the Cancer Outcomes and Services Dataset (COSD). Other fields are routinely derived from other data sources, such as Cancer Waiting Times, Cancer Screening Programmes and the Office for National Statistics (these include derived fields such as Charlson co-morbidity index and the Index of Multiple Deprivation (IMD) score). | N/A   | Jan 1985 - Dec 2015                   | Tumour level                  |
| Cancer registration (tumour table)  |  | N/A   | Jan 1985 - Dec 2015                   | Tumour level                  |
| Cancer registration (treatment table)   |  | N/A   | Jan 1985 - Dec 2015                   | Tumour level                  |
| <b>Other data sets for which Public Health England is the data controller</b> |  |   |                                       |                               |
| Route to diagnosis  | 'Route to Diagnosis' is defined as the sequence of interactions between the patient and the health care system, which lead to a diagnosis of cancer, based on the setting of diagnosis, the pathway and the referral route into secondary care. These derived fields robustly categorise the route to a cancer diagnosis into 8 categorical variables. The methodology for deriving these fields is described in detail in the British Journal of Cancer article " <a href="#">Routes to Diagnosis for cancer - Determining the patient journey using multiple routine data sets</a> ".  | Derived fields routinely linked to incident cancer cases from:<br>Jan 2006 – December 2013. | Jan 2006 – Dec 2013                   | Linked at a tumour level      |

|   |   |   |  |                                  |
|---|---|---|--|----------------------------------|
| Systemic Anti-Cancer Therapy Data set (SACT)  | <p>The Systemic Anti-Cancer Therapy data set collects clinical information on patients receiving cancer chemotherapy in or funded by the NHS in England. It relates to all cancer patients, both adult and paediatric, in acute inpatient, day case, outpatient settings and delivery in the community. SACT covers chemotherapy treatment for all solid tumour and haematological malignancies and those in clinical trials. The data set has been designed to collect information on all drug treatments with an anti-cancer effect, in all treatment settings, including traditional cytotoxic chemotherapy and all newer agents.</p>  | <p>Linked to incident cancer cases from:<br/>April 2012 – Dec 2015<br/><br/>(Data available for these patients up to Sept 2016)</p>                           | <p>April 2012 onwards<br/>(6 months behind current date)</p>   | <p>Linked at a patient level</p> |
| Radiotherapy Data Set (RTDS)  | <p>The Radiotherapy Data Set (RTDS) collates data from all NHS Acute Trust providers of radiotherapy services in England to provide consistent and comparable data on radiotherapy services in England. Data is collected for all radiotherapy of the following types:</p> <ul style="list-style-type: none"> <li>• Teletherapy</li> <li>• Brachytherapy given using automated remote afterloading machines</li> </ul> <p>All other brachytherapy given for the treatment of malignant disease delivered in England to patients in NHS facilities, or in private facilities where delivery is funded by the NHS, from 1st April 2009.</p>   | <p>Routinely linked to incident cancer cases from:<br/>April 2009 – Dec 2015<br/><br/>(Radiotherapy data for these patients are available up to Jan 2017)</p> | <p>April 2009 – onwards<br/>(3 months behind current date)</p> | <p>Linked at a patient level</p> |
| Quality of Life of Colorectal Cancer Survivors in England: Patient Reported Outcome Measures Survey (PROMS) | <p>The Quality of Life of Colorectal Cancer Survivors in England: Patient Reported Outcome Measures includes responses from respondents to a questionnaire distributed to colorectal cancer patients in January 2013, who were alive 12-36 months after a diagnosis of colorectal cancer. Respondents (n=34,467) described different parts of their journey from a diagnosis of colorectal cancer to treatment and through to their experiences of aftercare. The outcome questions in the survey are made up of three instruments: the EQ-5D (Euroqol 5 level), FACT items (Functional Assessment of Cancer Therapy) and SDI (Social Difficulties Inventory).</p> <p><a href="https://www.england.nhs.uk/wp-content/uploads/2015/03/colorectal-cancer-proms-report-140314.pdf">https://www.england.nhs.uk/wp-content/uploads/2015/03/colorectal-cancer-proms-report-140314.pdf</a></p> | <p>Routinely linked to incident cancer cases from:<br/>Jan 2010 – Dec 2011*</p>   | <p>Jan 2010 – Dec 2011</p>                                     | <p>Linked at a patient level</p> |

|   |  |  |                              |                                  |
|---|--|--|------------------------------|----------------------------------|
| <p>Quality of Life of Cancer Survivors in England: Pilot Patient Reported Outcome Measures Survey (2011) (Breast, Colorectal, Prostate, Non-Hodgkin's Lymphoma)</p> | <p>The Quality of Life of Cancer Survivors in England: Pilot Survey (2011) was commissioned by the Department of Health as part of the National Cancer Survivorship Initiative (NCSI). The survey was conducted by Quality Health in conjunction with cancer registries. The survey measures overall quality of life of representative samples of cancer survivors with four different tumour types (breast, colorectal and prostate cancer and non-Hodgkin's Lymphoma (NHL)) and at four different time points after diagnosis (approximately one, two, three or five years). Responses were obtained (n=3300) of individuals approach to participant, giving a response rate of 66%. Outcome questions in the survey are made up of three instruments: the EQ-5D (Euroqol 5 level), FACT (Functional Assessment of Cancer Therapy) items and SDI (Social Difficulties Inventory).</p> <p><a href="https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/267042/9284-TSO-2900701-PROMS-1.pdf">https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/267042/9284-TSO-2900701-PROMS-1.pdf</a></p> | <p>Routinely linked to incident cancer cases from:<br/>July 2006 - July 2010 *</p> | <p>July 2006 - July 2010</p> | <p>Linked at a patient level</p> |
| <p>Quality of Life of Cancer Survivors in England: Pilot Survey One Year Follow-Up (2012) (Breast, Colorectal, Prostate, Non-Hodgkin's Lymphoma)</p>                | <p>This is a resurvey of the cohort above; with an 85% response rate.</p>  | <p>Routinely linked to incident cancer cases from:<br/>July 2006 – July 2010*</p>  | <p>July 2006 – July 2010</p> | <p>Linked at a patient level</p> |
| <p>Patient Reported Outcome Measure Survey in Bladder Cancer: 2013 pilot</p>  | <p>The bladder cancer pilot PROMs survey was commissioned in 2012 with data collected January to March 2013. These surveys were undertaken by Picker Institute Europe on behalf of the Department of Health. The questionnaire was sent to 1,252 people who had been diagnosed with bladder cancer, and 673 completed questionnaires were returned, a response rate of 54%.</p> <p><a href="https://www.england.nhs.uk/wp-content/uploads/2015/10/proms-bladder-cancer.pdf">https://www.england.nhs.uk/wp-content/uploads/2015/10/proms-bladder-cancer.pdf</a></p>   | <p>Routinely linked to incident cancer cases from:<br/>Jan 2007 - March 2011*</p>  | <p>Jan 2007 - March 2011</p> | <p>Linked at a patient level</p> |

|   |  |   |   |                                  |
|---|--|---|---|----------------------------------|
| <p>Patient Reported Outcome Measures Survey in Gynaecological Cancers: 2014 pilot</p> | <p>This is a collection of three PROMs surveys conducted for cervical, ovarian and womb cancer. These surveys were undertaken by Picker Institute Europe on behalf of NHS England. Surveys were sent in 2013 to women with cervical, womb and ovarian cancer; most of whom had their initial treatment between 1 and 5 years previously. The questionnaire was sent to 1,252 women who had been diagnosed with cervical cancer, and 493 completed questionnaires were returned, a response rate of 39%. The questionnaire was sent to 1,252 women who had been diagnosed with ovarian cancer, and 685 completed questionnaires were returned, a response rate of 55%. The questionnaire was sent to 1,252 women who had been diagnosed with womb cancer, and 654 completed questionnaires were returned, a response rate of 52%. <a href="https://www.england.nhs.uk/cancer/resources/">https://www.england.nhs.uk/cancer/resources/</a></p> | <p>Routinely linked to incident cancer cases from: 2008 – 2012*</p>   | <p>2008 – 2012</p>  | <p>Linked at a patient level</p> |
| <p>National Audit of Cancer Diagnosis in Primary Care</p>                             | <p>The National Cancer Diagnosis Audit will collect primary care data from GP surgeries across England, Wales, Scotland and Northern Ireland for patients diagnosed with cancer in 2009/10. This data allows us to develop a picture of how cancer is diagnosed across the UK by looking at what symptoms patients present to the GP, how many and what tests or investigations patients are sent for and the time taken for a GP to refer patients to specialist care. Interested stakeholders can then use this data to influence policy and care changes to improve early diagnosis of cancer in the UK. The audit involved 1,170 practices and yielded data on 18,879 patients from 20 Cancer Networks.</p>  | <p>This cannot be linked to the cancer registry data</p>  | <p>April 2009 – March 2010 (plus some data outside this period)</p> | <p>Linkage not possible</p>      |
| <p>Lung Cancer Data Audit (LUCADA)</p>  | <p>The Lung Cancer Data Audit (LUCADA) looks at the care delivered during referral, diagnosis, treatment and outcomes for people diagnosed with lung cancer and mesothelioma. The data items in the LUCADA data set have been compiled to meet the requirements of the audit, and are not to be confused with the data items identified as Lung Cancer in the National Cancer data set. The audit focuses on measuring the care given to lung cancer patients from diagnosis to the primary treatment package, assessing</p>   | <p>The audit follows patients diagnosed between: Jan 2005 – Dec 2013<br/><br/>(The vital status of each patient, when linked to the registry data, can be</p> | <p>Jan 2005 – Dec 2013</p>  | <p>Linked at a patient level</p> |

|                                   |   |   |   |                           |
|-----------------------------------|---|---|---|---------------------------|
|                                   | <p>against standards and bringing about necessary improvements. The project supports the Calman Hine recommendations, the National Cancer Plan and other national guidance (e.g. NICE guidance) as it emerges.</p>  | followed up to Dec 2016)  |   |                           |
| National Lung Cancer Audit (NLCA) | <p>For the first time in England, the audit uses data collected and processed by the National Cancer Registration and Analysis Service (NCRAS). This replaces the previous bespoke dataset submitted by trusts through a web portal called LUCADA (Lung Cancer Audit Dataset). Alongside lung cancer multidisciplinary teams (MDTs) submitting data using the Cancer Outcomes and Services Dataset (COSD) to the NCRAS, the final dataset includes other registry and national datasets submitted by trusts, including pathology reports, Hospital Episode Statistics (HES), the National Radiotherapy Dataset (RTDS), the Systemic Anti-Cancer Therapy (SACT) dataset and death certificates. This linkage of many datasets provides the most comprehensive picture of lung cancer care to date.</p>   | Routinely linked to incident cases from Jan 2015 – Dec 2015   | Jan 2015 – Dec 2015                               | Linked at a patient level |
| National Prostate Cancer Audit    | <p>The aim of the National Prostate Cancer Audit (NPCA) is to assess the process and outcomes of prostate cancer care provided by the NHS in England and Wales. The audit collects data on:</p> <ul style="list-style-type: none"> <li>• The characteristics of all men with newly diagnosed prostate cancer, how their cancer was detected, and the referral pathway.</li> <li>• The crucial steps in the diagnostic and staging process.</li> <li>• The planning of initial treatment.</li> <li>• Initial treatments</li> <li>• Initial health outcomes.</li> </ul> <p>In addition, the NPCA collects patient-reported experience and outcome measures in the men included in the audit who underwent a radical treatment one year after diagnosis.</p> <p>The data set consists of a number of COSD data items with additional audit-specific items.</p> | <p>Routinely linked to incident cancer cases from:</p> <p>April 2014 – Dec 2014</p> <p>(The vital status of each patient, when linked to the registry data, can be followed up to Dec 2016)</p> | April 2014 onwards (3 months behind current date) | Linked at a patient level |

\*As PROMS data is a survey sent out at a specific point in time, the data provides a 'snapshot' of the patient's reported outcomes, so no other data is available outside of these diagnosis dates.

Table 2: Data sets accessible through the NCRAS service:

| Data set   | Description  | Years linked to the cancer registry   | Data controller  | Minimum level of linkage type    |
|--|--|---|--|----------------------------------|
| <p>HES admitted care</p> <p>HES outpatient</p> <p>HES accident and emergency</p> | <p>Hospital Episode Statistics (HES) are administrative data that detail admissions to NHS hospitals in England. The information is collected into three main datasets: Inpatient (including maternity), Outpatient and Accident &amp; Emergency. HES data contains data on all inpatient admissions, outpatient appointments and A&amp;E attendances at NHS hospitals in England and is collected during a patient's time at hospital. It is routinely submitted to NHS Digital to allow hospitals to be paid for the care they deliver.</p>  | <p>Routinely linked to incident cancer cases from:<br/>April 1997 - Dec 2014</p> <p>(Hospital episode data for these patients are available up to Dec 2015)</p>   | <p>Please contact the NHS Digital Data Access and Request Service (DARS)</p> | <p>Linked at a patient level</p> |
| <p>Diagnostic Imaging Data Set</p>   | <p>The Diagnostic Imaging Data Set (DID) is a central collection of detailed information about diagnostic imaging tests carried out on NHS patients, extracted from local radiology information systems and submitted monthly. The DID captures information about referral source and patient type, details of the test (type of test and body site), demographic information such as GP registered practice, patient postcode, ethnicity, gender and date of birth, plus items to calculate waiting times for each diagnostic imaging event, from time of test request through to time of reporting.</p>  | <p>Routinely linked to incident cancer cases from:<br/>1 April 2012 – Dec 2014</p> <p>(Imaging information for these patients are available up to July 2016)</p>  | <p>Please contact the NHS Digital Data Access and Request Service (DARS)</p> | <p>Linked at a patient level</p> |
| <p>National Cancer Waiting Times Monitoring Data</p>                             | <p>The National Cancer Waiting Times Monitoring Data set supports the continued management and monitoring of all the cancer waiting times standards. These include a maximum two week wait from urgent GP referral for a suspect cancer to the date first seen by a specialist; a maximum 31-day wait from diagnosis to first definitive treatment for all cancers; a maximum 62-day wait from referral from a Cancer Screening Programme to first definitive treatment. Other waiting times are included and NCRAS should be contacted for the entire list. As well as data specifically related to the waiting times, the data includes other useful information on treatments received etc.</p> | <p>Routinely linked to incident cancer cases from:<br/>Jan 2009 – Dec 2014</p> <p>(Cancer waiting times data for these patients are available up to Dec 2015)</p> | <p>Please contact NHS England</p>  | <p>Linked at a patient level</p> |

|  |  |  |   |   |
|--|--|--|---|---|
| National Cancer Patient Experience Survey (CPES) | The National Cancer Patient Experience Survey (CPES), commissioned by NHS England through Quality Health is a survey sent out to all adult cancer patients (aged 16 and over) with a primary diagnosis of cancer who have been admitted to an acute or specialist NHS Trust in England providing adult cancer services as inpatients or day cases, and discharged within a specified three month sampling period each year. The survey aims to collect information from patients about their cancer journey from their initial GP visit prior to diagnosis, through diagnosis and treatment and to the ongoing management of their cancer.   | <p><b>Wave 1:</b> 2010: patients discharged between 01/01/10 – 31/03/10</p> <p><b>Wave 2:</b> 2011/12: patients discharged between 01/09/11 – 30/11/11</p> <p><b>Wave 3:</b> 2013 patients discharged between 01/09/12 – 30/09/12</p> <p><b>Wave 4:</b> 2014 patients discharged between 01/09/13 – 30/11/13</p> | Please contact Quality Health   | 80% of surveys linked to a tumour level |
| Data for Head and Neck Oncology (DAHNO)          | The National Head and Neck Cancer Audit was commissioned and sponsored by the Healthcare Quality Improvement Partnership, HQIP developed in partnership with BAHNO (British Association of Head and Neck Oncologists) and managed by the NHS Digital. This national audit focuses on cancer sites within the head and neck (excluding tumours of the brain and thyroid cancers) the most common being the larynx and in the oral cavity. The audit works by collecting data from hospitals within England & Wales which diagnose and treat patients with cancer of the head & neck. Analysis is undertaken to measure individual trust and cancer network performance against the national IOG and BAHNO (British Association of Head & Neck Oncologists). These results are published in an annual report Open data available at: <a href="https://data.gov.uk/dataset/national-head-and-neck-cancer-audit-open-data-2014">https://data.gov.uk/dataset/national-head-and-neck-cancer-audit-open-data-2014</a> | <p>Routinely linked to incident cancer cases from:<br/>Nov 2008 – Oct 2013</p> <p>(The vital status of each patient , when linked to the registry data, can be followed up to Dec 2016)</p>  | Contact the Healthcare Quality Improvement Partnership (HQIP)         | Linked at a patient level               |
| NBOCAP   | <p>The National Bowel Cancer Audit (NBOCAP) was established to describe and compare the diagnosis, care and outcomes of patients diagnosed with bowel cancer.</p> <p>The Audit collects data on items which have been identified and generally accepted as measures of good care.</p>  | <p>The audit follows patients diagnosed between:<br/>April 2006 - March 2013</p> <p>(The vital status of each patient , when linked to the registry data, can be followed up to Dec 2016)</p>  | Please contact the NHS Digital Data Access and Request Service (DARS) | Linked at a patient level               |

First published: November 2016

© Crown copyright 2017

Re-use of Crown copyright material (excluding logos) is allowed under the terms of the Open Government Licence, visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/3/](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/) for terms and conditions.