NDRS
NATIONAL DISEASE REGISTRATION SERVICE

NHS Digital

Cancer registration methodology

About the NDRS

The National Disease Registration Service (NDRS) is part of NHS Digital (NHSD). Its purpose is to collect high-quality, timely data on cancer, rare diseases and congenital anomalies to monitor changes in the health of the population.

The NDRS includes:
- The National Cancer Registration and Analysis Service (NCRAS); and
- The National Congenital Anomaly and Rare Disease Registration Service (NCARDRS).

Healthcare professionals, researchers and policy makers use data to better understand population health and disease. The data is provided by patients and collected by the NHS as part of their care and support. The NDRS uses the data to help:

- understand cancer, rare diseases and congenital anomalies;
- improve diagnosis;
- plan NHS services;
- improve treatment;
- evaluate policy;
- improve genetic counselling.

NDRS
NATIONAL DISEASE REGISTRATION SERVICE

National Disease Registration Service
NHS Digital (NHSD)
The Leeds Government Hub
7&8 Wellington Place
Leeds
LS1 4AP

For queries relating to this document, please contact:
NDRSenquiries@nhs.net

Improving lives with data and technology – NHS Digital support NHS staff at work, help people get the best care, and use the nation's health data to drive research and transform services.

NHS Digital

Contents

Introduction

This document describes the methodology for the Cancer registration statistics for England publication produced by the National Cancer Registration and Analysis Service (NCRAS) within NHS Digital (NHSD).

The Cancer registration statistics for England is an annual publication that reports on the number and rate of new cancer registrations and deaths due to cancer in England and includes sub-national geographies.

The numbers and rates are broken down by cancer site, socio-economic status, sex and age group for all malignancies. For selected cancers, the number and rate of new registrations by stage at diagnosis and Index of Multiple Deprivation (IMD) are also reported.

The data used in all these publications is part of the National Cancer Registration Dataset , data collected are subject to quality assurance and control described in the methodology on Data collection and Quality assurance of administrative data for cancer registrations.

To increase the timeliness of the cancer registration statistics, NCRAS may choose to publish two releases. The first release, based on provisional data, will contain a summary of counts and rates of cancer diagnoses for the registration year for England only using appropriate groups of diagnosis codes.

The full analysis of cancer registrations including stage at diagnosis, based on finalised data, will then be published within another six months, and will also contain data for regions of England. The data will be presented for all diagnosis codes, grouped and ungrouped.

If the time to having the finalised data is within three months, or there is a significant delay to registering provisional data the full analysis only may be published. There may be minor differences in estimates provided between the first and final release that arise from the finalisation process.

All publications consist of data tables and an accompanying bulletin. The data tables are produced in OpenDocument Spreadsheet format (.ods) and the bulletin is produced in html format to ensure accessibility. For the diagnosis year 2019 onwards, the cancer registrations can be found NHS Digital Cancer Registration Statistics and previous publications are on gov.uk cancer registration statistics collection page.

Methodology

All cancers are coded using the International Classification of Disease 10th Revision (ICD-10). ICD-10 coding for cancer is based on the nature and anatomical site of the cancer. In the cancer registration statistics publication, the term “invasive cancer” is used to refer to all invasive tumours (C00 to C97 in ICD-10 coding), excluding non-melanoma skin cancers (NMSC) (ICD-10 C44).

NMSC are common cancers, we report only the first tumour per person for both basal cell carcinomas and cutaneous squamous carcinomas and therefore this represents an under-estimation of the total tumour count.

Time series for cancer registrations are published with and without NMSC. This provides users with consistent and meaningful statistics over time.

The cancer registration statistics publication also report on selected benign tumours (ICD-10 D32 to D33 and D35.2 to D35.4), all in situ tumours (D00 to D09) and tumours with uncertain or unknown behaviour (D37 to D48).

Cancer Incidence

To complete the Cancer registration statistics for England publication, NCRAS takes a snapshot of the cancer data from the Cancer Analysis System when the expected registrations for a year are complete. This provides a single, consistent source of cancer registrations.

The snapshot allows NCRAS to publish consistent statistics on the:

- Counts of registrations of cancer by cancer type (3- and 4-digit ICD-10 codes), age, sex and deprivation (from 2019 registration year)
- Age-specific rates of cancer incidence by cancer type (3- and 4-digit ICD-10 codes) and sex
- Age-standardised and non-standardised rates of cancer incidence by cancer type (3- and 4-digit ICD-10 codes), sex and deprivation (from 2019 registration year)

From the 2018 registration year, the percentage of cancer registrations by stage at diagnosis for all stageable common cancers which have at least 70% completeness are also reported in the Cancer registration statistics publication. The data reports diagnoses from 2013 onwards.

TNM is the staging system used for most cancer sites and has 4 main stages.
The International Federation of Gynaecology and Obstetrics (FIGO) system is used to stage gynaecological cancers, the Ann Arbor system or Murphy St Jude system are used to stage Hodgkin lymphomas, the International Neuroblastoma Risk Group Staging System (INRGSS) is used to stage neuroblastoma, the Chang's staging system is used to stage medulloblastoma; these all have 4 broad stages. The International Staging System (ISS) is used to stage myelomas, the Binet staging system is used to stage chronic lymphocytic leukaemia; these both have 3 stages.

NCRAS uses the TNM to complement the staging information provided by the other systems, except for cervix cancers where only FIGO is used. This is because the previous version of FIGO for cervix did not include nodal status (N component of TNM). The new version of FIGO for cervix does now include nodal status, but is slightly different to current TNM staging.

Stage at diagnosis of cancer is an important factor that can affect treatment options offered to patients and outcomes, including survival. Earlier diagnosis, that is, usually when cancers are diagnosed at stages 1 and 2 as opposed to stages 3 and 4, is associated with better prognosis on average. The government aims to have 75% of the top 10 cancers diagnosed at stages 1 and 2 by 2028.

It should be noted that not all cancers have a staging system, for example, most brain cancers do not currently have a staging system. This is why estimates of incidence by stage at diagnosis cannot be provided for all types of cancer.

The rates presented in the Cancer registration statistics publication are known as cancer incidence rates or cancer mortality rates. A cancer incidence rate is the number of new cancer registrations made in a year within a population. A cancer mortality rate is the number of deaths caused by a cancer in a year within a population.

Age-specific rates are rates by age-group; incidence and mortality rates are usually given per 100,000 population.

Incidence rate = (number of cancer registrations/population) x 100,000

Age-standardised rates are a weighted average of the age-specific cancer rates. All age-standardised cancer rates published since June 2014 have used the 2013 European Standard Population (ESP). The ESP is a set of weights for a population structure, based on an average European age-profile of a population. The weights are applied to age-specific rates to produce age-standardised rates.

Age-standardised incidence rates allow users to make fair comparisons of cancer diagnoses:

- across different regions in England
- with other countries in the UK or Europe
- between different time periods

Note: the ESP has an upper age band of 95 and older. England population estimates produced by Office of National Statistics (ONS) currently have an upper age band of 90 and older. This means NCRAS must combine the weights for the 90-94 and 95 and older age bands. The upper age band for incidence or mortality calculations is also 90 and older.

Cancer Mortality

ONS provides NCRAS with extracts from their annual mortality registration files. ONS document their dataset in the mortality data quality and methodology report. NCRAS uses the dataset to publish the statistics for:

- Number of deaths from cancer by cancer type (3-digit ICD-10 codes), age, sex, and deprivation (from 2019 registrations onwards)
- Age-specific rates of cancer mortality by cancer type (3-digit ICD-10 codes) and sex
- Age-standardised and non-standardised rates of cancer mortality by cancer type (3-digit ICD-10 codes), sex and deprivation (from 2019 registrations onwards)

All rates are expressed per 100,000 population and age-standardised rates are standardised to the 2013 ESP.

Statistical accuracy

Rates calculated on low numbers can be hard to interpret. To help users make reliable interpretations, the data reference tables for the publication include:

- Age-specific rates where the count for a combination of age, sex and cancer site is 3 or more.

- Directly age-standardised rates where the total count for a combination of age and cancer site is 10 or more

To help users make reliable interpretations, the publication tables DO NOT include:

- Age-specific rates where the count for a combination of age, sex and cancer site is fewer than 2 or less
- Directly age-standardised rates where the total count for a combination of age and cancer site is 9 or less

The publication tables suppress rates with “u” if they are not included for the reasons stated above. Additionally, the publication tables will flag any non-standardised or directly age-standardised rates based on a total count of between 10 and 19 with a “u”. This is a warning to users that the small number of diagnoses may affect the reliability of these rates.