

Information paper

Quality and Methodology Information

General details

Title of output: Cancer survival by stage at diagnosis
Designation: Experimental Statistics
Geographic Coverage: England
Date of last SQR or QMI* N/A
Contact details: cancer.newport@ons.gov.uk

Executive summary

ONS publishes a suite of bulletins on cancer incidence and survival, with accompanying Quality and Methodology Information (QMI) documents for cancer [incidence](#) and cancer [survival](#). This Cancer Survival by Stage at Diagnosis QMI relates to the new experimental statistics for England on Cancer Survival by Stage at Diagnosis, which has been produced jointly with Public Health England (PHE). Survival estimates have been produced by stage of diagnosis for adult English patients with cancer (15-99 years). The tumour groups included are: bladder (ICD code C67), female breast (C50), colorectal (C18-C20), kidney (C64), lung (C33 to C34), melanoma (C43), ovary (C56 to 57), prostate (C61) and uterus (C54 to C55).

This paper provides details of the methodology used to create these experimental statistics. It has been prepared jointly with the National Cancer Registration and Analysis Service of PHE.

Cancer is a major cause of death, accounting for around a quarter of deaths in England. More than 1 in 3 people will develop cancer at some point in their life. In January 2011, the Government published [Improving Outcomes – a Strategy for Cancer](#). This document sets out how the Government plans to improve cancer outcomes, including improving survival rates through reductions in the proportion of patients who are diagnosed with cancer at an advanced stage. Diagnosis at non-advanced stage is associated with a notably improved probability of surviving cancer. Specific public health interventions, such as screening programmes and information/education campaigns aim to decrease the proportion of cancer patients diagnosed at advanced stage. The tumour sites listed above represent the most commonly occurring stageable tumour sites and are also included in the [Public Health Outcomes Framework](#) (PHOF) experimental indicator 2.19 (Cancer diagnosed at an early stage). We have excluded non-Hodgkin lymphoma because between patients with the same stage survival varies substantially depending on tumour sub-types.

This document contains the following sections:

- Output quality
- About the output
- How the output is created
- Validation and quality assurance
- Concepts and definitions
- Other information, relating to quality trade-offs and user needs
- Sources for further information and advice

* Quality and Methodology Information (QMI) reports replaced Summary Quality Reports (SQR) from April 2011

Output quality

This document describes the quality of the output and details the points that should be noted when using the output.

ONS has developed [Guidelines for Measuring Statistical Quality](#). These are based upon the 5 European Statistical System (ESS) Dimensions of Quality and several other important quality characteristics:

- Relevance
- Timeliness and punctuality
- Coherence and comparability
- Accuracy
- Output-quality trade-offs
- Assessment of user needs and perceptions
- Accessibility and clarity

More information is provided about these quality dimensions in the sections below.

About the output

Relevance

(The degree to which the statistical outputs meet users' needs.)

For a number of years there has been an increasing interest in cancer survival by stage. Cancer survival is generally influenced by stage at diagnosis, treatment quality and host factors (for example frailty or comorbidity). When comparing the survival of all patients with cancer in different countries it is therefore difficult to know how much of this variation relates to differences in diagnosis, treatment or host factors. In contrast examining variation in survival by stage can elucidate the likely role of factors *other than* stage at diagnosis on cancer outcomes.

The importance of reporting cancer statistics by stage at diagnosis had been emphasised in the ONS output [consultation in 2012](#). With improvements in the collection of cancer registrations by PHE over the last few years it is now possible to create cancer survival estimates by stage at diagnosis for the most common stageable tumour sites. The definition of cancer stage is agreed internationally by professional bodies. This bulletin is the first time that such statistics are being made available across England as part of routine reporting. The analysis has been released as experimental in the first instance to allow us to gather views and opinions on the analysis undertaken.

This experimental series presents 1-year age-standardised net cancer survival for tumours diagnosed in England during the three years 2012, 2013 and 2014 followed up to the end of 2015, with an overall estimate of survival from all malignant cancers and 9 common cancers; bladder, breast (female only), colorectal, kidney, lung, melanoma, ovary, prostate and uterus.

Important users of cancer survival estimates include the Department of Health, academics and researchers, cancer charities, cancer registries, other government organisations, researchers within ONS, the media and the general public. The Department of Health uses cancer survival estimates to brief government ministers, and as part of the evidence base to inform cancer policy and programmes, for example in drives to improve cancer survival.

Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

From the registration of 2012 cancer diagnoses onwards, the use of the same data collection tools and methodologies across England has enabled a consistent approach to collecting and recording cancer staging data. This new approach to registering tumours successfully captured more than 60% of all stageable tumours being staged for the first time in 2012. For 2013 and 2014 diagnoses, the proportion staged has further improved to 71% and 77% respectively. To allow trends in survival by stage to be monitored, this publication will be updated annually as future cohorts of diagnostic data are made available for analysis.

For more details on related releases, the [GOV.UK website](#) provides 12 months' advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will

be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Official Statistics](#).

How the output is created

General principles of the methodology for this output

The National Cancer Registration and Analysis Service holds cancer registration data in PHE from (then) regional cancer registries covering the entire population of England since the 1960s. Each year, more than 200,000 patients are newly diagnosed with cancer in England. The [Health and Social Care Information Centre \(HSCIC\)](#) routinely updates these individual cancer records with information on each patient's vital status (alive, emigrated, dead or not traced). At the time that data were extracted for the most recent series of statistical bulletins, less than 0.5% of patients diagnosed were not traced during the relevant period.

Survival estimates are age-standardised to improve the comparability between population groups and over time. This is because net survival varies with age at diagnosis, and the age profile of cancer patients can vary over time and between geographical areas.

The estimates produced in this publication have been produced by the National Cancer Registration and Analysis Service of PHE. These experimental statistics implements the [United Kingdom and Ireland Association of Cancer Registries'](#) ratified Standard Operating Procedure [Guidelines on Population Based Cancer Survival Analysis](#) and we are using a cohort approach combined with the [Pohar-Perme net survival estimator](#), as implemented by Paul Dickman & Enzo Coviello's [strs](#) programme in Stata v13.1. The analysis in the bulletin presents 1 year net age-standardised cancer survival calculated using a cohort approach. A 1 year survival estimate has been produced for patients diagnosed 2012 - 2014. The estimates have been calculated using the International Cancer Survival Standard age-weightings. The life tables used were supplied by the [London School of Hygiene and Tropical Medicine's Cancer Survival Group](#). These English life tables for 2011 were used for 2012 to 2014. In these life tables, the mortality of cancer patients is compared to that of individuals in the general population who belong to the same group of age, sex, socio-economic status and region.

The following [criteria](#) are used to identify the patients that are eligible to be included in the analysis (and provided in the inclusion table included as part of the publication release):

- patients should have a unique identifier
- patients should have a complete date of birth and be aged between 15 and 99 at diagnosis
- patients who have died should have a complete registered date of death
- patients should have a complete date of cancer diagnosis
- patients should have a known sex
- patients should have a known date of being recorded as alive or dead
- patients should be resident in England and have a valid postcode for usual place of residence
- tumours should be malignant and be newly diagnosed in the studied cohort and be a primary tumour
- cancers of the blood (lymphomas, leukaemias & myelomas) should not occur in a solid tumour
- patients are not excluded if they had further primary tumours of the cancer of interest later in the period of interest, nor if they had any primary tumour of another cancer site diagnosed in the period of interest, nor if they had any type of primary tumour diagnosed before or after the period of interest
- patients should not be Death Certificate Only (DCO) registrations where the first confirmed record of a tumour occurs after the death of the patient
- the sequence of dates should be valid (for example a patient should not be diagnosed before they are born)

Other decisions applied include:

- where more than 1 primary tumour of interest is diagnosed in the same period for a patient, the first date of diagnosis is used to calculate the period of survival experienced by that patient
- where a patient dies on the date of diagnosis but is not a DCO registration, then these patients should be included in the survival analyses but should have 1 day added to the recorded date of death to accommodate limits in Stata's [stset](#) command. Survival estimates were generated at monthly intervals up to 6 months following diagnosis, then at 3-monthly intervals up to 1 year following diagnosis

- the weights outlined by the [International Cancer Survival Standard](#) are used in the age-standardisation

Structured Query Language (SQL) was used to extract data from NCRAS Cancer Analysis System and Stata code used for preparing the data and calculating these estimates. To aid transparency and allow replication of the estimates produced in this bulletin, annotated copies of the SQL and Stata code can be provided on request (NCINenquiries@phe.gov.uk). Data can be applied for through the Office of Data Release at PHE (ODR@PHE.gov.uk).

Validation and quality assurance

Accuracy

(The degree of closeness between an estimate and the true value.)

Cancer survival releases are produced using the most robust methods available for population-based cancer survival estimation, using documented quality assurance procedures. Both the size of the cohort and survival rates were compared to previous (relevant published^{1, 2} and unpublished) estimates and found that the estimates were similar to those published in [Cancer survival in England by stage](#) after making due allowances for the differences in methodology and selection criteria.

Cancer incidence data for England are collected by the regional offices of the National Cancer Registration and Analysis Service, which is part of [Public Health England](#). Data are submitted to the National Cancer Registration and Analysis Service from a range of healthcare providers and other services (for example, pathology laboratories). The quality and accuracy of the data submitted by different sources may vary. The regional offices of the National Cancer Registration and Analysis Service collate all the data for each patient, including checks for internal consistency of the sequence of dates, as well as the cancer site, sex, morphology and duplicate registrations. These checks are closely based on those published by the [International Agency for Research on Cancer \(IARC\)](#).

If a record fails any critical validation check - for example if the date of birth is invalid - the records are not reported in Cancer Registration Statistics, or any other ONS publication, including survival releases, since it is not possible to send these records for verification of the patient's vital status to the [Health and Social Care Information Centre \(HSCIC\)](#). If a record passes all critical validation checks, or fails 1 or more minor quality controls, these records are sent to the HSCIC for verification of vital status.

Further checks are required for survival analysis; these are carried out in 2 stages. The first stage involves checking the eligibility of a record based on its completeness, the patient's usual residence, tumour behaviour and morphology. Patients with an invasive, primary, malignant tumour are eligible for analysis (see section How the output was created which lists the full criteria). Ineligible patients include those whose tumour is benign (not malignant) or *in situ* (malignant, but not invasive) or of uncertain behaviour, or for which the organ of origin was unknown.

The second stage involves checking the patient's age (15-99 years for adults), vital status, that the patient's sex is compatible with the cancer site, the dates are valid, and the patient was not registered solely from a death certificate.

Since 2001, cancer registrations for each year have been estimated to be between 96-99% complete at the time of extraction, with completeness improving over time. However, it is important to note that the cancer registration database is dynamic. In common with cancer registries in other countries, cancer incidence in England can take up to 5 years after the end of a given calendar year to reach 100% completeness and stability, because of late registrations, corrections and deletions. The figure of 99% for the 2012 data year is an estimate of completeness based on the figures published for the 3 previous years, compared with the number of late registrations subsequently received for these years. It is therefore the difference between the figures published in [Cancer Registration Statistics](#) (and all subsequent ONS cancer incidence publications within that reporting year), and late registrations received after the publication date cut-off. It is not an estimate of how many cancers are never recorded.

Coherence and comparability

(Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain for example, geographic level.)

This bulletin is the first time that cancer survival by stage statistics is being made available across England as part of routine reporting. It is therefore not possible to directly compare the stage-specific survival estimates directly to other England cancer survival outputs. Care should be taken if the “all stage” estimate in this bulletin is compared to previous national statistics for 1-year site specific cancer survival estimates as there can be differences. This is caused by differences in the computation methods, for example PHE use the International Cancer Survival Standard weights while the weights from Coleman *et al*² are used in the production of the National Statistics. The issue of comparability of cancer survival statistics across the UK is being discussed at the [UK and Ireland Association of Cancer Registrations \(UKIACR\)](#) Executive Board.

International comparisons of cancer survival figures are occasionally reported. Care should be taken when interpreting results from different countries, because of known differences in healthcare and cancer registration systems, which are likely to affect on results. A discussion of the issues raised by comparison of survival figures from different countries was published as part of the [International Cancer Benchmarking Partnership \(ICBP\) Study](#). Other international sources of survival by stage statistics include those published by [National Cancer Institute](#) in the USA and the [Canadian Cancer Society](#).

Cancer survival estimates are published at England level by various organisations and they will not all be directly comparable. Raw data may be taken from different sources and differences in quality assurance procedures will influence final estimates.

[Cancer Research UK](#) publishes relative cancer survival estimates by geography, deprivation level, cancer site, age at diagnosis and sex.

The [Northern Ireland](#), [Scottish](#) and [Welsh](#) registries publish national figures for their respective countries. The period for which most recent data are available may differ between countries. Where possible, ONS cancer survival bulletins present data for more than 1 time period, to enable comparison.

Concepts and definitions

(Concepts and definitions describe the legislation governing the output, and a description of the classifications used in the output.)

Cancer

Definition - Cancers are coded using the [International Statistical Classification of Diseases Tenth Revision \(ICD-10\)](#). ICD-10 coding for cancer is based on the nature and anatomical site of the cancer. ICD-10 replaced the previous revision, ICD-9, in 1995.

For the purposes of the national cancer registration scheme, the term “cancer” includes all malignant neoplasms (tumours that invade into surrounding tissues), which are conditions listed under site code numbers C00 to C97 in ICD-10. In addition, all *in situ* (malignant but not invasive) neoplasms (D00–D09), certain benign (not malignant) neoplasms (D32–D33, D35.2–D35.4), and neoplasms of uncertain or unknown behaviour (uncertain whether benign or malignant, D37–D48) are registered.

Primary cancer

Definition - A primary cancer is the tumour that first develops in an identifiable part of the body, for example the stomach, and usually gives the name to the type of cancer with which a patient is diagnosed. A primary tumour that has been found growing in, or spread to, another part of the body is called a secondary tumour or metastatic tumour. A spread of primary tumour cells within the system of lymph nodes is not usually considered to be metastases of the primary tumour.

Cancer stages

Definition - Cancer stages at diagnosis are a measure of how far the primary tumour has grown when the patient first presents in hospital and are staged according to internationally agreed standards. The most common staging standards are agreed by the Union for International Cancer Control (sometimes called the TNM staging method) and stage according to three components:

tumour size (the T component), nodal involvement of the lymphatic system (N) and metastatic spread (M). Some gynaecological cancers are staged using an alternative method set out by the International Federation of Gynaecology and Obstetrics (FIGO). For cancers of the ovary and the uterus, FIGO stages can be uniquely matched to TNM stages and this has been used to supplement the TNM staging data.

Although the combinations of tumour size, nodal involvement and metastatic spread change by tumour type, generally there are four stages of cancer progression:

- Stage 1: The primary tumour is usually small and is contained within the body organ in which the tumour started growing
- Stage 2: Although larger, the primary tumour has not spread to other parts of the body. Spread to the lymphatic system may be included depending on the primary tumour site
- Stage 3: The primary tumour is larger and may have spread into neighbouring parts of the body and into the lymphatic system
- Stage 4: The primary tumour has spread to at least 1 other part of the body, creating a secondary or metastatic tumour

Some samples taken do not produce clear results and some patients are too unwell to undergo the surgery required to obtain a tissue sample for staging. There are several reasons why a tumour cannot be staged, for example if the tumour size, nodal involvement or metastatic involvement is not recorded for a particular tumour then it may not be possible to derive an overall stage at diagnosis. In this analysis missing stage was treated as a separate category and survival estimates produced for patients with 'missing stage at diagnosis' alongside the other categories of known stage.

Definitions

Net survival - the survival of cancer patients after taking into account the background mortality that patients would have experienced if they had not had cancer. The Pohar-Perme estimator of net survival is an unbiased estimator that accounts for "informative censoring" bias, that is the fact that some groups of patients are less likely to be observed for the full duration of follow-up than others, independently of their cancer prognosis.

Age-standardisation - a technique used to enable metrics derived for different populations to be robustly compared even when the populations are different in terms of their age profiles. Age-standardised rates also allow for more robust comparisons between males and females, years, and geographical areas.

95% confidence interval - a measure of the uncertainty around an estimate. Confidence intervals provide a range of values which contain the true population parameter with a 95% level of confidence.

Cohort approach - when follow-up information is available for each patient for at least 1 year, 1-year survival can be estimated using the classical cohort approach. For example, once follow-up information is available for each patient until 31 December 2015, the cohort approach can be used to estimate 1-year survival for patients diagnosed during 2014, or earlier, by combining the conditional probabilities of survival to the end of each successive sub-period of the analysis.

Legislation

The Statistics and Registration Service Act 2007 permits the Registrar General to provide to the United Kingdom Statistics Authority, in order to carry out any of its functions, both information that is kept under the Births and Deaths Registration Act 1953 and any other information received by the Registrar General in relation to any birth or death.

The Health Service (Control of Patient Information) Regulations 2002 Statutory Instrument No. 1438, Regulation 2, permits confidential patient information relating to patients referred for the diagnosis or treatment of cancer to be processed for the following purposes:

- the surveillance and analysis of health and disease
- the monitoring and audit of health and health related care provision and outcomes where such provision has been made
- the planning and administration of the provision made for health and health related care

- medical research approved by research ethics committees the provision of information about individuals who have suffered from a particular disease or condition where:
 - that information supports an analysis of the risk of developing that disease or condition
 - it is required for the counselling and support of a person who is concerned about the risk of developing that disease or condition

This regulation was made under Section 60 of the Health and Social Care Act 2001 and continues to have effect under Section 251 of the NHS Act 2006.

ONS processes and stores cancer registration data in accordance with the requirements of:

- the Data Protection Act 1998
- the Code of Practice for Official Statistics
- the IACR/IARC Guidelines on Confidentiality in Cancer Registries (Coleman MP, Muir CS, Ménégos F. Confidentiality in the cancer registry. Br J Cancer 1992; 66: 1138-49)
- the European Network of Cancer Registries (ENCR) Guidelines on Confidentiality (2002)
- the United Kingdom Association of Cancer Registries (UKACR) Guidelines on Confidentiality 2012
- the NHS Act 2006
- the Statistics and Registration Service Act 2007

PHE processes and stores cancer registration data in accordance with the requirements of:

- The Data Protection Act 1998
- The Health Service (Control of Patient Information) Regulations 2002
- The Caldicott Report 1997
- Information: To share or not to share? The Information Governance Review March 2013
- The Freedom of Information Act 2000
- Section 251 of the NHS Act 2006 (originally enacted under Section 60 of the Health and Social Care Act 2001)
- Confidentiality: NHS Code of Practice 2003
- NHS Records Management Code of Practice (Part 1, 2006 & Part 2, 2009)
- Health and Social Care Act 2012
- The NHS Information Security Management Code of Practice 2007
- The Computer Misuse Act 1990
- The Human Rights Act 1998

Other information

Assessment of user needs and perceptions

(The processes for finding out about uses and users, and their views on the statistical products.)

A stakeholder review of all ONS cancer publications was conducted in 2010. Stakeholders were asked for their views about how they use the relevant outputs, their importance and their quality. Comments were also sought on any changes respondents would like to see in terms of content and presentation of the outputs and of the ONS cancer web pages. The results of this [consultation](#) can be found on the [ONS website](#).

A stakeholder [consultation](#) of all ONS cancer publications was undertaken in 2012 to determine future user needs. The results of this consultation can be found on the [ONS website](#). One of the key needs identified as part of this consultation was for data on stage at cancer diagnosis, which are collected under the National Cancer Registration Scheme and collated by PHE. Due to improvements made by PHE in the collection of stage information as part of cancer registration, PHE in partnership with ONS are now able to publish estimates of survival by stage diagnosis of cancer in the form of experimental statistics. ONS will continue working with PHE to ensure that such survival estimates will be published as National Statistics in the future.

We welcome feedback from users on the content, format and relevance of our statistics. Please contact cancer.newport@ons.gsi.gov.uk.

Sources for further information or advice

Accessibility and clarity

(Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

For information regarding conditions of access to outputs, please refer to the links below:

- [Terms and conditions \(for data on the website\)](#)
- [Accessibility](#)

In addition to this Quality and Methodology Information document, quality information relevant to each release is available in the background notes of each of the statistical bulletins, and in the metadata contained within the downloadable tables.

All cancer survival statistical bulletins are web-only releases, available in either HTML or PDF formats; data tables are available in Excel format. For further information about cancer survival bulletins, please contact:

Cancer Analysis Team (Room D165)
Office for National Statistics
Government Buildings Cardiff Road
Newport
NP10 8XG
Telephone: 01633 455704

Email: cancer.newport@ons.gsi.gov.uk

Useful links

[UK and Ireland Association of Cancer Registries](#)
[National Cancer Registration and Analysis Service](#)

References

1. Quaresma, M, Coleman, MP, Rachet, B, 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971-2011: a population based study, *The Lancet* (2015) DOI: 10.1016/s0140-6736(14)61396-9 www.thelancet.com
2. S. McPhail et al., Stage at diagnosis and early mortality from cancer in England *British Journal of Cancer* (2015) 112, S108–S115. doi:10.1038/bjc.2015.49 www.bjcancer.com
Published online 3 March 2015
3. Coleman MP, Babb P, Damiecki P, Grosclaude PC, Honjo S, Jones J, et al. Cancer survival trends in England and Wales 1971-1995: deprivation and NHS Region. (Studies on Medical and Population Subjects No. 61). London: The Stationery Office; 1999. p.55-106.