



Macmillan-NCIN UK Cancer Prevalence Project

Guidance and FAQs June 2015 releases

Topics

[Background of the Macmillan-NCIN work plan](#)

[The UK Cancer Prevalence Project](#)

[What is cancer prevalence?](#)

[Why is cancer prevalence important?](#)

[Cancer prevalence – why are the numbers rising?](#)

[What timeframe do these cancer prevalence statistics cover?](#)

[Why does the prevalence data only go up to 2010?](#)

[All cancers combined prevalence counting method](#)

[Cancer site prevalence counting method](#)

[Analysis for multiple primaries available in the UK summary table](#)

[The different counting methods used](#)

[The exclusion criteria](#)

[What cancer sites are covered?](#)

[What geographical areas are covered?](#)

[What is the difference between “age at diagnosis” and “age at the end of 2010”?](#)

[Disclosure of small numbers – rounding method](#)

[How does this prevalence data compare to other prevalence data available?](#)

[Why are deprivation scores for the UK combined not shown?](#)

[What if patients left the country after diagnosis?](#)

[Cancer site and topic issues to consider](#)

[Acknowledgements](#)

[References](#)

Background of the Macmillan-NCIN work plan

The Macmillan-NCIN work plan is a joint initiative between Macmillan Cancer Support and the National Cancer Intelligence Network (NCIN) to improve the understanding of the 2.5 million people currently living with cancer in the UK¹. We want to know more about cancer survivors, where they are and the issues they face, so that future services can be better tailored to each person's needs and improve their health outcomes. You can read more on the Macmillan-NCIN work plan [here](#).

The Macmillan-NCIN UK Cancer Prevalence Project

The Macmillan-NCIN UK Cancer Prevalence Project aims to quantify and characterise the UK cancer population so that there is a more granular understanding of them. The datasets provide new cancer prevalence statistics for people diagnosed with cancer over a 20 year period (1991-2010), that is 20-year prevalence.

What is cancer prevalence?

We use the term 'cancer prevalence' for the number of people who have been diagnosed with cancer in the past and who are still alive, on a given date (the index date). The index date we have used here is 31 December 2010. The 20-year period we cover is 1991-2010. Data for Northern Ireland are only available from 1993 when the registry began collecting data.

We use 'cancer survivors' to mean anyone who has been diagnosed with cancer in the given period, regardless of whether their cancer has been cured, is in remission or is still present.

Why is cancer prevalence important?

Cancer prevalence statistics are essential for understanding the complete burden of cancer i.e. the number of people who have previously received a diagnosis of cancer and their need for care. Some people will be receiving active treatment for their cancer, some may be recovering from cancer and readjusting to life after treatment and others may be living with the on-going consequences of their cancer and its treatment.

Cancer prevalence – why are the numbers rising?

Our 20-year cancer prevalence statistics act as an indicator of the burden of cancer and can help to inform health and social care service planning. As well as the number of people getting cancer, the number of people surviving cancer is also increasing. The number of people living with cancer is influenced by new diagnoses and cancer survival main factors include:

- A growing and ageing population: life expectancy is increasing, with more elderly people alive today than ever before so there are more people being diagnosed with cancer.
- Increasing cancer survival: thanks to a greater focus on early diagnosis and advances in cancer treatments.
- Changing lifestyle risk factors: for example increases in obesity rates, and decreases in smoking prevalence.

What timeframe do these cancer prevalence statistics cover?

We have analysed cancer registrations data for a 20-year period, 1991-2010, for England, Scotland and Wales, and for an 18-year period, 1993 to 2010, for Northern Ireland (NI). Northern Ireland started recording cancer registration data from 1993 onwards, hence a shorter period was used for this nation. More up to date data and 20-year prevalence are now available directly from the Northern Ireland Cancer Registry [here](#). Data published elsewhere may not be directly comparable with the other UK nations' data.

As some people would have been diagnosed with cancer before 1991 (or 1993 for NI) and are still alive, our current prevalence figures are based on '**Limited Duration Cancer Prevalence**'. As such, the overall figure for UK cancer prevalence presented in this spreadsheet is around 1.8 million based on the last 20 years.

'Complete Cancer Prevalence' is based on all people who had a diagnosis of cancer and are still alive, on a given date. It has been estimated that around "2.5 million" people in the UK are still alive after having had a diagnosis of cancer¹. The next stage of our study aims to estimate the complete number of people living with cancer beyond 20 years since diagnosis.

Why does the prevalence data only go up to 2010?

At the time of data extraction (the year 2013), the latest data available in the National Cancer Data Repository (NCDR) was for 2010. Cancer incidence and mortality data are subject to a number of quality assurance processes before they are available for national analysis. So there will often be a gap of 2-3 years between the year of data collection and release for analytical purposes.

Since this time, the English centralised cancer registration system (ENCORE) began to feed cancer data into a central Public Health England database (Cancer Analysis System). This will be useful for the next phase of this project.

All cancers combined prevalence 'first specific' counting method

Figures for all malignant neoplasms combined are based on the first cancer diagnosed in each person, within the period 1991-2010. This approach aligns with the methods used to estimate the "2.5 million" people living with a cancer diagnosis in the UK¹.

The figures for all cancers combined relate to all malignant cancers registered, excluding non-melanoma skin cancer (NMSC) (ICD10 codes C00-C97, excluding C44). We routinely exclude NMSC in cancer data analyses because the condition is often diagnosed and treated within primary care services, such as GP surgeries, and is not consistently recorded as a cancer diagnosis across UK countries.

Cancer site prevalence counting method

Across all tables, except the UK summary table, the cancer site data is based on first diagnosis of each specific cancer - for example, if a person was diagnosed with lung cancer in 1991 and colorectal cancer in 2005, that person would be included in counts for both lung and colorectal cancers. If that person was diagnosed with lung cancer in 1991 then with lung cancer again in 1997, that person would be counted only once in lung cancer, based on their first lung cancer diagnosis in 1991.

The cancer site statistics relate to tumours diagnosed, rather than people. As such some people will be double-counted if they have been counted with two or more cancers of a different site, within the 20-year period.

Work is ongoing to scale up these 20-year prevalence figures to **complete prevalence**, i.e. the total count of people who are alive and have ever had a cancer diagnosis, and update figures and methods first published in 2009².

Analysis for multiple primaries available in the UK summary table

The [UK Summary Table](#) is a stand-alone table that shows the cancer site prevalence figures using a "[first only](#)" counting method, as well as the "[cancer site prevalence counting method](#)" outlined above. As such, the difference between the two sets of

numbers shows us how many tumours were diagnosed in people who had received a previous diagnosis of cancer:

Example for lung cancer:

<i>First diagnosis only method (as denoted above)</i>	<i>Cancer site first-specific method (as denoted above)</i>
Number of people: 52,759	Number of tumours: 57,178
Difference = 4,419	

If we use the example above, we can see that almost 8% of lung tumours registered were in people previously diagnosed with a cancer of a different site, within the 20-year period. This is around 4,400 people, or 1 in 13 people living with lung cancer, who previously received a cancer diagnosis of a different site in the period 1991–2010.

Please note that the data cannot tell us what the chronology of all cancer diagnoses were, nor at what sites they occurred.

For the next stage of the UK Cancer Prevalence Project we will investigate multiple primary tumours in more detail. This will involve further analysis of the cancer registrations data, which are based on a longer period of time (not just the 20 years). We will explore all second and subsequent cancers (not just the second and subsequent diagnoses of a different cancer site). We will also explore further the findings and the context of these data by discussing the information with experts, comparing our findings with other studies and having a deeper understanding of cancer registration practice.

The different counting methods used

We analysed the cancer prevalence data using the different counting methods mentioned earlier, after consultation with internal and external stakeholders.

It is not possible to use the all cancers combined numbers as a base denominator with the cancer site figures. The all cancers combined figures have been extracted using a different method than the cancer site figures, so these are not directly comparable, or to be used together.

The exclusion criteria

If a person died or left the country in the given period (1991-2010), or were aged over 99 at diagnosis or over 105 at the end of 2010, they were removed from the study.

What cancer sites are covered?

We present cancer prevalence data for each UK nation and the UK combined by cancer group, 21 common cancer sites and 47 detailed cancer sites. There is a worksheet in each data file, to show the ICD-10 codes used for each cancer site. We engaged with a number of stakeholders to decide on the cancer sites to be included.

What geographical areas are covered?

We present cancer prevalence data for each UK nation and the UK combined by cancer group, 21 common cancer sites and 47 detailed cancer sites. For cancer groups and common cancer sites, the data is segmented by age at end of 2010, sex and time since diagnosis. An additional worksheet also shows common cancer sites presented by age at diagnosis, sex and time since diagnosis.

We present cancer prevalence data by the following sub-national areas for all ages combined, by sex, time since diagnoses and 21 common cancer sites:

- England – by Local Area Team, CCG, Local Authority, Strategic Clinical Network
- Wales – by Cancer Network, Health Board, Local Authority and upper super output area
- Scotland – by Cancer Network, NHS Board Area and Local Council Area
- Northern Ireland – by Health and Social Care Trust and Local Government District

The aforementioned geographies were current at the time of data extraction in 2013/14, so they do not take into account any boundary changes since this time.

What is the difference in the analysis by “age at diagnosis” and “age at the end of 2010”?

We have provided cancer prevalence data by both age at diagnosis and age at end of 2010. Both sets of figures provide us with valuable information. By looking at age at diagnosis we can look at how old people were at the time of their diagnosis and then how long they have been living with cancer.

The analysis by age at the end of 2010 is useful in showing us the attained age of people that have previously been diagnosed with cancer, to show the current ages of cancer survivors at the end of the period.

What does living with cancer mean for a patient and why is time since diagnosis helpful?

The cancer prevalence numbers are defined as the number of people who are currently alive and have had a cancer diagnosis in the past - in this case for people diagnosed in the 20-year period 1991-2010. For some of these people the cancer will be influencing most aspects of their lives, whilst for others the impact will be less acute.

The cancer prevalence data are provided for the total 20-year period and by periods of time after first diagnosis: 0-1 years, 1-2 years, 2-5 years, 5-10 years, 10-15 years and 15-20 years.

Segmenting the prevalent cancer population in this way provides an insight into how many people have recently been diagnosed with cancer, and how many are longer term survivors. The total prevalence figures for the 20-year period are a useful indicator of the burden of cancer as a whole, however, patient needs and experiences will vary over time after diagnosis. Segmenting the prevalent population by time since diagnosis can therefore help to inform health care service planning for the needs of people living with cancer at different time periods on their pathway.

The time since diagnosis (TSD) categories relate to the following years of diagnosis:

TSD	Diagnosed in...
0-1 yrs	2010
1-2 yrs	2009
2-5 yrs	2006-2008
5-10 yrs	2001-2005
10-15 yrs	1996-2000
15-20 yrs	1991-1995

Disclosure of small numbers – rounding method

As the data are segmented by a number of variables, meaning the numbers can get very small, we had to find a suitable method to remove the risk of disclosure of potentially identifiable patient data.

The following geographical areas required suppression, which was applied according to the requirements of each UK nation's cancer registry:

- Northern Ireland: all areas, including nation-level for values <5;
- England: Area Team, Clinical Commissioning Group and Local Authority for values <6;
- Scotland: All sub-national areas <6;
- Wales: USOA-level only <6.

In order to preserve the maximum amount of values, we applied a rounding suppression method to the values for the areas mentioned above. For all values 0-2, values were rounded down to zero; all values 3-4 were rounded up to 5 for Northern Ireland, and values 3-5 up to 6 for all other UK nations.

Whilst this method was the optimum choice for preserving the maximum number of values and granularity, this does mean that the figures within our datasets sometimes differ. We recommend that when you extract figures from the datasets, to use for your own purposes, that you use the following footnote:

Some of the Macmillan-NCIN 20-year cancer prevalence statistics use a rounding suppression method. These statistics are accurate enough to represent the number of tumours diagnosed in cancer survivors, but do note that there may be minor differences with other sources.

How does this prevalence data compare to other prevalence data available?

The numbers in this analysis may not agree with those published elsewhere due to slight differences in methodologies, data extraction methods, periods of observation, datasets, and rounding.

Why are deprivation scores for the UK combined not shown?

Deprivation analyses have been provided, though they are country-specific. Each UK nation has a different index of deprivation and, as such, we are unable to combine them.

What if patients left the country after diagnosis?

We removed all patients and their registered tumours from our dataset, if the patients had been flagged as leaving the country of diagnosis. There may be patients who moved from one UK nation to another, but their tumour registrations will not be captured. So if someone was diagnosed with cancer when resident in England, then later moved to Wales within the period (and this embarkation had been captured in the cancer database), they would have been removed from the prevalence dataset.

Cancer site and topic issues to consider

There have been variations in the collection and registration practices across the UK cancer registries. For instance, bladder figures will include some uncertain behaviour (D41.4) and in-situ bladders (D09.0) pre-2007. Other registries implemented this coding change earlier than Wales, hence the higher number of cases in Wales compared to other UK countries.

It is recommended that childhood cancers be registered using the [International Classification of Childhood Cancers](#) (ICCC) system, and for teenagers and young adults the Birch classification system³ should be used. We were not able to extract the cancer prevalence cohort by morphological tumour type at the time of extraction, so we were limited to ICD-10 classification only. It is suggested that you take this into consideration when analysing the cancer prevalence data for the age groups 0-14 and 15-24.

We have included some benign and in-situ D-codes in the cancer site figures. (As standard, these are not included at all in the all cancers combined figures (see [All cancers combined prevalence counting method](#) for more information)). For example, for brain and central nervous system (CNS) tumours there are a number of benign D-codes included – these are particularly important for brain/CNS tumours, as even benign tumours can be life-threatening in these areas of the body, and are therefore clinically significant.

Please note that Northern Ireland cervix in-situ figures (D06) are not included in our 20-year prevalence data.

We have provided the data in hidden sheets within each workbook, so that these may be unhidden and used for your own analyses.

Any questions?

If you wish to have support in interpreting the results, or have any comments, you can contact ncinanalysts@phe.gov.uk.

If you would like to know more about the Macmillan-NCIN partnership you can also contact evidence@macmillan.org.uk.

Acknowledgements

Working together

The Macmillan-NCIN UK Cancer Prevalence Project is part of the Macmillan Cancer Support and Public Health England's National Cancer Intelligence Network Partnership Work Plan. The Project is supported by the Knowledge & Intelligence

Team (East) at Public Health England. Data are sourced and presented in collaboration with the Welsh Cancer Intelligence and Surveillance Unit, Health Intelligence Division, Public Health Wales, the Scottish Cancer Registry and the Northern Ireland Cancer Registry, funded by the Public Health Agency for Northern Ireland.

[Full acknowledgements and summary of key contributors](#)



References

1. Maddams J, Utley M, Møller H. Projections of cancer prevalence in the United Kingdom, 2010–2040. *Br J Cancer* 2012; 107: 1195–1202. (Projections scenario 1). Macmillan analysis based on extrapolation of 2010 and 2020 projections that the number of people living with cancer will hit an estimated 2.5 million in 2015.
2. J Maddams, D Brewster, A Gavin, et al. Cancer Prevalence in the United Kingdom: estimates for 2008. *Br J Cancer* 2009; 101(3):541–547.
3. Birch JM, Alston RD, Kelsey AM, et al. Classification and incidence of cancers in adolescents and young adults in England 1979-1997. *Br J Cancer* 2002;87:1267-74.