MACMILLAN CANCER SUPPORT

Overview of the Available UK-Wide Cancer Prevalence Outputs

The following document aims to provide a summary of the sources and available outputs of cancer prevalence figures for each individual nation and the UK, including the pros and cons of consulting one dataset over another.

When considering cancer prevalence, it's important to stress the difference between complete and observed prevalence data:

- **Current complete, or total, modelled prevalence** represents the total number of people ever diagnosed with cancer who are still alive at a specific time point, also known as the index date.

  Because cancer registries won't have population-level data for the years prior to their establishment, estimating complete cancer prevalence involves statistical modelling applied to observed prevalence to make up for past missing cancer registrations.

  Although complete prevalence figures are impactful in communications, its reliance on statistical modelling, rather than recorded data, means that it is rarely possible to calculate estimates for local areas and rare cancers, whereas it is possible to report on these within observed prevalence.

- **Future predicted complete prevalence** is an estimate of complete prevalence calculated by projecting changes to the prevalent population (estimated through complete prevalence methods) into the future. This can be achieved by modelling future trends for the incidence and survival of cancer types in different age groups.[i]

- **Observed, or limited duration, prevalence** refers to the number of people diagnosed with cancer during a specific time period who are still alive on the index date. This limited duration, normally 20 years or more, is dependent on how long a registry has been collecting population-level data on cancer registrations. This makes observed prevalence data a more accurate representation of a population living with cancer over a specific period of time. In addition, it is also possible to report on the observed prevalence for local areas and rare cancer types. For this reason, observed prevalence data is preferable for any statistical work where accurate and granular information are priorities.

Whatever the cancer prevalence data of choice, it's important never to mix complete and observed prevalence data in any calculations or report unless a clear explanation of the different methodologies is included.

[i] Maddams J, Utley M, Møller H. Projections of cancer prevalence in the United Kingdom, 2010-2040. Br J Cancer 2012; 107: 1195-1202.

MACMILLAN CANCER SUPPORT

These are the main complete and observed prevalence data sources for UK-wide cancer prevalence:

Complete prevalence:

- **Macmillan's predicted complete prevalence in 2020, based on Maddams et al's modelling.**
  This estimate is an update for 2020 of Macmillan's 2.5 million figure of complete cancer prevalence in the UK. It is based on the following published measures of observed cancer prevalence for each nation, in the UK:

  - **England:** Cancer Prevalence in England. 23-year prevalence (1995-2017), by Transforming Cancer Services Team for London (TCST), NHS and National Cancer Registration and Analysis Service (NCRAS), PHE.

  - **Northern Ireland:** 2017 Cancer prevalence data. 25-year prevalence (1993-2017), by Northern Ireland Cancer Registry (NICR), Queen's University Belfast.

  - **Scotland:** Prevalence of Cancer in Scotland, 20-year prevalence (1998-2017), by Scottish Cancer Registry, Information Services Division (ISD).

  - **Wales:** Local Cancer Intelligence (LCI) Wales. 21-year prevalence (1995-2015), provided by Welsh Cancer Intelligence and Surveillance Unit (WCISU) and published by Macmillan Cancer Support.

  The observed prevalence figures for people living with cancer for the longest period of time since diagnosis available (20 to 23 years for England, 15 to 25 for Northern Ireland, 10 to 20 for Scotland and 10 to 21 for Wales) were extracted from each of these observed prevalence datasets above.

  The next step was to calculate the ratios of people living with cancer for these observed time periods (20 to 23 years for England, 15 to 25 for Northern Ireland etc.) to the remaining population living with cancer (between 20 years for England, 15 years for Northern Ireland, 10 years for Scotland, 10 years for Wales to a maximum of 66 years since diagnosis) for each cancer type, sex and nation. We did this using complete prevalence data for 2013 from the Macmillan-NCRAS Cancer Prevalence Project.

  These ratios were used to scale up the whole population. For example, the England prevalence scaled up the observed prevalence for people living with cancer 20 to 23 years since diagnosis to an estimation of the prevalence for people living with cancer 20 years to 66 years since diagnosis. This estimate was then added to the observed prevalence for patients living with cancer 0-19 years since diagnosis to estimate complete cancer prevalence.

  This process was carried out for all cancers combined (excluding non-melanoma skin cancer), and for female breast, prostate, colorectal and lung cancer separately.

For each of these estimates of complete cancer prevalence, further predictive estimates were then produced for 2020 and following years. These predictions were based on UK growth rates in the prevalence projections from Maddams et al (2012) (based on scenario 1), using an exponential curve matching known data points. Growth rates were applied to individual cancer types and genders for both 0 to 5 years since diagnosis and all years of time since diagnosis.

Pros:

- Currently, this figure is the most contemporary estimate available for complete cancer prevalence.
- Although it's technically a figure of predicted prevalence, it is closely aligned with the outputs of the Macmillan-NCRAS UK Cancer Prevalence Project.
- The estimate is planned to be widely quoted in Macmillan's external communications, with the intention of it replacing the previous 2.5 million as an established figure for the number of people currently living with cancer, in the UK.

Cons:

- The observed prevalence figures, which form the basis of these calculations, are only currently available up to 2015 or 2017, as there is a delay in their compilation due to the process of cancer registration.
- These estimates of complete prevalence rely on multiple assumptions, including that the UK growth rates predicted by Maddams et al (2012) are still relevant and apply to each nation equally. The growth estimates are likely to be less accurate for predictions further into the future from the observed prevalence data on which they are based. The Macmillan predicted complete prevalence figures are therefore heavily rounded to acknowledge this.

When to use:

- Macmillan's updated figure for complete cancer prevalence in 2020 is the best estimate available for any nationwide communications related to the number of people currently living with cancer in the UK. The availability of projections of this estimate, based on Maddams et al (2012), to 2025, 2030 and 2040 can also be used to demonstrate the estimated growth of the cancer population.

Where to find:

- To be published at macmillan.org.uk/cancerprevalence in Q1 2020.

- **Phase 3 of the Macmillan-NCRAS UK Cancer Prevalence Project**
  In September 2016, Phase 3 of the UK Cancer Prevalence Project between Macmillan and NCRAS produced a dataset of complete cancer prevalence for the UK and its nations, for all and the top four cancers, with an index date of 2013. Unlike Macmillan's predicted complete cancer prevalence projections, Phase 3's calculations aren't projecting future trends in cancer prevalence, and are therefore the **most accurate figures of UK-wide complete cancer prevalence** to date. This project modelled the complete prevalence in the UK in 2013 as 2.27 million. Due to its less recent index date of 2013, and an increase of 2.27 million in 2013 to 2.5 million in 2015 being credible based

on the projections from Maddams et al (2012), Macmillan made the decision to continue to quote the previously mentioned 2.5 million figure for people living with cancer in the UK in 2015. However, the output from Phase 3 of the UK Cancer Prevalence Project has been extremely valuable in validating the 2.5 million complete prevalence projection figures, as well as forming the basis of scaling up observed to complete prevalence in the new Macmillan complete prevalence projections for 2020.

Pros:

- As the calculations of complete cancer prevalence are based on observed and modelled prevalence, rather than projections, their accuracy is likely to be higher, and the numbers much closer to the actual number of people living with cancer in the UK in 2013.
- The workbook contains splits of the data by time since diagnosis, which provide information on when people initially received a cancer diagnosed.

Cons:

- The dataset's index date of 2013 makes it less appealing for headline statistics as these aim to focus on reporting the numbers closest to current prevalence.
- Figures from both the Phase 3 complete prevalence project (2.27 million) and the 2.5 million complete prevalence in 2015 projections can only be further broken down by the top four cancers, which limits their use in work focused on specific, less common cancer types, as well as health boards and local authorities across the UK. For these more granular analyses, observed prevalence is the best data source.

When to use:

- We recommend the complete cancer prevalence dataset from Phase 3 of the UK Cancer Prevalence Project for statistical projects where accuracy of the figures takes priority over being able to quote a more recent index date.

Where to find:

- UK Complete Cancer Prevalence for 2013 Workbook for UK nations and UK combined.
- UK Complete Cancer Prevalence for 2013 Technical Report describing the methodology used to calculate complete cancer prevalence estimates.
- A paper describing the work is also under development.

Observed prevalence:

- **National releases of observed cancer prevalence**
  These datasets provide records of the number of people diagnosed and surviving cancer within an observed number of years, and still alive at a specified index date, for each nation. In contrast to complete prevalence, observed prevalence presents **more accuracy with greater granularity**, based solely on registrations without modelling, including data on many cancer types, age at diagnosis and time since diagnosis.

  However, the available figures are limited by how long the relevant cancer registry has been established. Currently, the most recent dataset of observed cancer prevalence for all England covers a 23-year period, with an index date of 2017. This was produced as a collaboration between Transforming Cancer Services Team for London (TCST), NHS and NCRAS, PHE. Cancer prevalence for Scotland and Northern Ireland has been released by Information Services Division (ISD) Scotland and the N. Ireland Cancer Registry (NICR). These examine observed periods of 20 and 25 years with an index date of 2017. For Wales, in partnership with the Welsh Cancer Intelligence and Surveillance Unit (WCISU), 21-year observed prevalence data with an index date of 2015 is published on Macmillan Cancer Support's Local Cancer Intelligence (LCI) tool, in the form of interactive tables and graphs.

  The Phase 3 Cancer Prevalence Project results can be used to quantify the difference between observed (limited duration) and complete (total) prevalence figures. Overall in the UK, this project estimated that 13% of patients had a diagnosis >20 years prior to the index date of 2013. This estimate varied between 12% (England) and 23% (NI) for different nations, and between 1% (prostate) and 18% (all other malignant neoplasms) for different cancer sites.

  Pros:
  - Measures of observed cancer prevalence offer the most accurate and up-to-date snapshots of the population living with cancer within each nation over a defined period of time.
  - Observed cancer prevalence presents granular information, including breakdowns by several cancer types and age at diagnosis, which are not available for complete prevalence measures.

  Cons:
  - The most obvious drawback of using observed prevalence in contrast to complete cancer prevalence is that it only covers periods of data between 20 and 25 years, depending on the national release considered. It therefore misses a number of people diagnosed with cancer before these dates and surviving to the index date. The size of the difference has been estimated as approximately 13% for 20-year limited duration prevalence in the UK, but varies depending on the nation and cancer site.

When to use:

- When considering cancer types other than the four most common (female breast, prostate, colorectal and lung), observed cancer prevalence datasets are the only sources available for prevalence.
- Where the priority is to quote the most recent figure available, based on actual records, for the number of people living with cancer in a specific nation.

Where to find:

- Cancer Prevalence in England in 2017: 23-year prevalence by demographic measures at Cancer Alliance, CCG, Upper and Lower tier Local Authority, and London Cancer Prevalence Dashboard 2017.
- Number and percentage of cancer survivors (prevalence) at 31 December 2017, Scotland.
- Cancer Prevalence for Northern Ireland: 1993-2017.
- Local Cancer Intelligence (LCI) Wales.

For more detailed information on prevalence methodologies see the Standard Operating Procedure: Guidelines for calculation of cancer prevalence.

MACMILLAN CANCER SUPPORT

Table of comparison of available cancer prevalence data for all cancers (ex. NMSC):

Type of prevalence measure	Duration	Index Date	England	Scotland	N. Ireland	Wales	UK	Cancers	Additional splits in the data	When to use
Complete, predicted[1]		2020	2,400,000	250,000	82,000	140,000	2,900,000	All combined and top 4	Sex, Time since diagnosis (0-5 & 5+ years).	Communications related to the estimated number of people living with cancer in the UK that require the most recent estimates
Complete[2]		2013	1,869,300	205,500	69,600	128,800	2,273,200	All combined and top 4		Statistical projects, in which the accuracy of the figures provided would take priority over being able to quote a more recent index date.
Limited Duration[3]	23-year	2017	1,953,647	-	-	-	-	All combined and many common sites	Age, Sex, Time since diagnosis, Local geographies.	When considering cancer types other than the four most common. When quoting based on actual records, for the number of people living with cancer in a specific nation.
Limited Duration[4]	20-year	2017	-	190,942	-	-	-			
Limited Duration[5]	25-year	2017	-	-	63,413	-	-			
Limited Duration[6]	21-year	2015	-	-	-	115,774	-			

[1] Macmillan's complete prevalence projection
[2] Phase 3 Macmillan-NCRAS UK Prevalence Project
[3] TCST NHS and NCRAS PHE
[4] ISD Scotland
[5] Northern Ireland Cancer Registry
[6] Welsh Cancer Intelligence & Surveillance Unit