April 2018

Standard Operating Procedure:

Guidelines for calculation of cancer prevalence

This SOP has been developed by the UKIACR Analysis Sub-Group, with important contributions from James Charnock (lead), Sarah Miller, Dyfed Huws, Anna Gavin, David Brewster, Vivian Mak, Joanna Pethick, Samuel Jones and Henrik Moller. The Analysis Group will maintain and update this document.

Further information on the content of this SOP is available from your regional or national cancer/public health lead. Their contact details are available from the UKIACR website (http://www.ukiacr.org/content/links).

Contents

1. Aim

There are a variety of different methods and techniques for calculating cancer prevalence across the UK nations and Ireland, depending on the purpose, data availability and detail required. The aim of this brief Standard Operating Procedure (SOP) is to present the various ways of calculating cancer prevalence and to encourage a consistent approach across the United Kingdom (UK) and Ireland.

With an increased appetite to undertake more UK and Ireland based analysis, having a consistent approach to calculating the cancer population to be able to compare across nations is all the more important and this SOP aims to address this gap in documentation.

This public access document is intended to be used by cancer and public health analysts involved in the analysis and calculation of cancer prevalence in the UK and Ireland, but also internationally. It is hoped that it will also be of wider interest to those tasked with compiling, understanding and interpreting the characteristics of the cancer population.

2. Introduction

Understanding the size and characteristics of the cancer population at a point in time is important for national health services. Without this information services are unable to adequately distribute healthcare services for the cancer population; this could potentially lead to poorer outcomes for cancer patients.

It is well documented that generally cancer incidence is rising[1] and, with the improvements in cancer survival[1] and an ageing population, the cancer population will only keep growing. Understanding how many people are living with and beyond cancer is essential to be able to provide the best care possible for this group of patients. As cancer prevalence is a composite measure of incidence and survival, it is a useful statistic by which to understand the impact of cancer on patients and health care services over the longer term.

Cancer prevalence is defined as the number of people diagnosed with cancer during a set period and who are still alive on a specified date; this differs from incidence, which observes the number of people diagnosed with cancer during a set time period, usually in one year. With regards to cancer, prevalence can be calculated in two main ways:

1) Person prevalence: the number of people diagnosed with cancer during a set period, in a defined population, and who are still alive on a specified date.
2) Tumour prevalence: the number of tumours diagnosed during a set time period in a defined population, and who are still alive on a specified date.

Each of these methods can also be sub-divided into two further categories:

1) Limited duration, or observed prevalence; the number of people diagnosed with cancer or tumours diagnosed during a set time period who are still alive at a designated end point. This end point is often referred to as the index or censor date.

2) Complete prevalence: the total number of people ever diagnosed with cancer or all tumours ever diagnosed who are alive at a designated end point.

Routine cancer registration and mortality records, such as those collected in the UK nations and Ireland, are important and allow us to calculate limited-duration and complete cancer prevalence. However, unless a disease registry has recorded every case of cancer for the period of interest and hold the corresponding patient mortality information; complete cancer prevalence will involve statistical modelling to account for years where data are not available.

As part of the Macmillan-NCRAS Work Plan, Macmillan Cancer Support have worked in partnership with Public Health England (PHE) and in collaboration with the cancer registries of Wales, Scotland and Northern Ireland, to develop three analytical projects based on generating cancer prevalence statistics, these are described below.

2.1. 20-year observed UK cancer prevalence

(i) The first phase of the UK Cancer Prevalence Project[2] aimed to provide an understanding of the cancer population and calculated cancer prevalence for people diagnosed with cancer between 1991 and 2010 and who were still alive on 31st December 2010*. This information is provided for all cancers combined and segmented by sex, age at diagnosis, age at end of 2010 and time since diagnosis, for the UK and each constituent nation.

(ii) The second phase of the project aimed to provide a more granular understanding of the UK cancer survivorship population, and calculated cancer prevalence for people diagnosed with cancer between 1991 and 2010 and who were still alive on 31st December 2010*. This was available by up to 47 site-specific cancers for the UK and each constituent nation, with additional breakdowns available by sex, age at diagnosis, age at end of 2010, time since diagnosis and deprivation quintile. Data was also provided for each nation at various sub-national levels for 21 most common cancers segmented by sex and time since diagnosis.

Both of these two analytical projects are examples of limited duration cancer prevalence.

* Except For Northern Ireland, which is based on patients diagnosed between 1993 and 2010 who were alive at the end of 2010 (18-year prevalence), due to the data available from the Northern Ireland Cancer Registry who commenced data collection in 1993.

Data for this project can be found at the following link:

http://www.ncin.org.uk/about_ncin/segmentation

This has now been updated for England through a partnership between NCRAS, Macmillan Cancer Support and the Transforming Cancer Services Team (London) (TCST) to provide 21-year limited duration detailed prevalence data.

This data can be found at the following link:

http://www.ncin.org.uk/local_cancer_intelligence/tcst

2.2. UK complete prevalence for cancer

(iii) The third phase of the UK Cancer Prevalence Project aimed to provide an estimate of complete cancer prevalence. This involved calculating estimates of the total number of people living with cancer in the UK who had *ever* been diagnosed with cancer and who were still alive at the end of 2013. Data for the UK and each constituent nation was calculated for all cancers combined (excluding non-melanoma skin cancer), female breast, prostate, colorectal and lung cancers, in addition to a group of all other cancers. This was segmented by year of diagnosis, sex and age at diagnosis.

This new work builds on previous research by Maddams et al. (2009)[3], in which it was estimated that there were 2 million people living with cancer in the UK at the end of 2008 and further estimating that 2.3 million people were set to be living with cancer at the end of 2013. This also showed alignment with previous analysis by Macmillan Cancer Support (based on Maddams et al. 2012)[4] in which it was estimated there would be 2.5 million people living with a cancer diagnosis in the UK by 2015.

Prevalence data (where available) for the individual nations is also available through each of the nation's websites found below:

Northern Ireland - http://www.qub.ac.uk/research-centres/nicr/

Wales - http://www.wcisu.wales.nhs.uk/home

Scotland - http://www.isdscotland.org/Health-Topics/Cancer/Scottish-Cancer-Registry/

3. Data preparation for prevalence calculations

3.1. Data quality

A standard quality control procedure should ensure that tumour records meet basic criteria of data quality, to include de-deduplication, consistent tumour site/morphology/behaviour/patient sex. The International Association of Cancer Registries (IACR) has the 'Check and Conversion Programs for Cancer Registries' software that runs internal validation and consistency checks on tumour records (see: http://www.iacr.com.fr under 'Support for registries'). This can be used to ensure all records pass the check software before beginning the analysis, also to make sure that all tumours are using the same coding system.

The overall prevalence statistics will be inflated if information about the death of cancer patients are not updated on the registration records, which may render them 'lost to follow-up'. To ensure we are using the correct follow-up for all patients, it is necessary to check that all death notifications have been received, processed and recorded. This is more relevant in registries that do not perform a complete linkage of their cancer incidence register to a national death register on a regular basis.

3.2. Data availability

Both limited duration cancer prevalence and complete cancer prevalence statistics are influenced by the availability of cancer incidence and general mortality data for the cancer population in each of the UK nations. The following describes the years of diagnosis of cancer registration data that are available for each UK nation:

England – 1971-2015

Scotland – 1971-2015

Wales – 1985-2015

Northern Ireland – 1993-2015

Data is also available for Ireland; however, to date work has only been carried out for the UK. Please contact the Ireland cancer registry for information regarding the data availability.

Data may be available for earlier dates; however, the data availability described here has been used in previous work. Please contact the individual registries for information on the availability of data.

4. Methods for calculating observed cancer prevalence statistics

4.1. Introduction

The following methodologies will focus on person prevalence. An explanation of how to calculate tumour prevalence will be provided at the end of this section.

For the example here, we will use England data for the diagnosis period 1971-2015. This will contribute to 45-year cancer prevalence data. However, this can vary by country based on the data availability mentioned above.

Cancer prevalence statistics can be calculated for any period of time. Five-year cancer prevalence figures will provide you with an awareness of more recent cancer survivors.

4.2. Inclusion and exclusion Criteria

A variety of criteria exist to determine if a cancer record should be included in a cancer prevalence cohort. Ineligible records should be recorded and verified in order to document data quality, and then excluded from cancer prevalence calculations, if appropriate. Here is a list of the commonly applied inclusion and exclusion criteria, used when generating cancer prevalence statistics for England.

4.2.1. Inclusion Criteria

1. Age at diagnosis – 0-99 years
2. Age at censor date <105 years. This is chosen to exclude patients that may have been lost to follow-up and would therefore artificially inflate the prevalence total
3. Tumour topography code (as defined by either International Classification of Diseases (ICD) or International Classification of Diseases for Oncology (ICDO) belongs to the appropriate definition of cancer site. Variations in coding definitions can affect prevalence calculations, especially when generating prevalence figures by cancer site. Only patients with primary tumours are included, i.e. tumours that have originated in the organ of the cancer site defined, and are not spread from another organ in the body (ie excluding secondary, recurrent or metastatic tumours).
4. Only patients with tumours of an invasive, primary, and malignant behavioural code (=3) in ICD 0 (http://www.who.int/classifications/icd/adaptations/oncology/en/) are included. It is worth noting that in revisions of ICD O the behavioural codes have changed. In ovarian cancer some behavioural codes 3 were reclassified as 1 (uncertain behaviour) when moving from the 2nd to 3rd edition of ICD 0; this has relevance when comparing historic estimates to recent estimates.
5. Sex = Male or Female

4.2.2. Exclusion Criteria

1. Subsequent primary tumour registrations should be excluded. As we are calculating person prevalence we are only concerned with the first tumour registered to a patient. If you are calculating tumour prevalence then this exclusion criteria is not relevant.
2. Exclude patient if the following information is missing or not a standard code (e.g. not listed in an official data dictionary): sex, date of diagnosis, date of birth or age.
3. Patients who are not resident in England (or in the nation of interest) at time of diagnosis. All patients dying on or before the censor date (for example, 31st December 2015). As we are interested in cancer prevalence, our cohort must be alive at the censor date to be included in the calculation. Any patient with a date of death after the censor date *should* be included in the analysis and their vital status should be set as alive for this purpose of this work.
4. Any patient with an embarkation date. Patients who have left the country, sent to prison or joined the armed forces at any point during the study should be removed from the analysis, as there is no guarantee that we can ensure their vital status is correct at the time of analysis.

4.3. Stratification Criteria

Once the cohort has been defined the stratification criteria for the cancer prevalence statistics need to be decided. The following categories are suggestions – they are dependent on how much of a geographical breakdown you would like. Using more granular sub-national breakdowns may affect the granularity of the other variables of interest.

- Geography

    - UK and national level
    - Sub-national e.g. Clinical Commissioning Group (CCG), Cancer Alliance or Sustainability and Transformation (STP) area in England. For other UK nations, this may be Health and Social Care Trust in Northern Ireland, Local Health Board in Wales or Health Board and NHS Region in Scotland.
- Age groups
    - The age of the patient could be displayed by age at diagnosis or age at the end of the follow up period. This must be decided before commencing the cancer prevalence analyses. ([1]example SQL code has been provided for calculating age at end of follow up), for example:
        - 5-year age bands (0-4, 5-9……. 95-99)
            - This is the lowest suggested age group to use, as this will allow combinations of age groups if required.
        - Age groups from previous Macmillan-NCRAS work
            - 0-14, 15-24, 25-44, 45-64, 65-69, 70-74, 75+
- Tumour Group
    - Please see suggested tumour groupings in the appendix
- Sex
    - 1 = males, 2 = females
- Deprivation
    - Due to data availability in each nation it is advised to categorise patients based on their deprivation at diagnosis.
        - For England, deprivation is calculated using the income domain from the indices of multiple deprivation.

4.4. Data disclosure control methods

Previous work has used the following criteria to adhere to national guidelines of data disclosure.

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.

As a general rule, suppression should occur when the underlying population of the area being used is below 1000 people however, please note that these are only guidelines and all suppression should

[1] TRUNC(TRUNC(MONTHS_BETWEEN(TO_DATE('31/12/2015','DD/MM/YYYY',DIAGNOSISDATE))/12))

be checked with each nation before publication. This also applies for final signoff of the release of this data.

4.5. Tumour prevalence

The only change that needs to be made for calculating tumour prevalence rather than person prevalence is to include subsequent cancers diagnosed per patient. This allows us to count all tumours. Note that patients may be counted more than once.

5. Methods for calculating complete prevalence

5.1. Introduction

The following methodologies will focus on person prevalence. An explanation of how to calculate tumour prevalence will be given at the end of this section.

For the purpose of the rest of this section we will be using England data with a range of 1971-2015. This will give us 45-year prevalence data, however, this can selected for any time series and any country using the data availability mentioned above.

5.2. Inclusion and exclusion Criteria

The inclusion and exclusion criteria listed below are identical to those in the previous section, however, for the sake of clarity they have been included here.

5.2.1. Inclusion Criteria

1. Age at diagnosis – 0-99 years.
2. Age at censor date <105 years. This is chosen to remove patients that may have been lost to follow-up and would inflate the prevalence total
3. Tumour topography code (as defined by either International Classification of Diseases (ICD) or International Classification of Diseases for Oncology (ICD0) belongs to the appropriate definition of cancer site. Variations in coding definitions can affect prevalence calculations especially when tabulating prevalence figures by cancer site. Only patients with primary tumours are included, i.e. tumours that have originated in the organ of the cancer site defined, and are not spread from another organ in the body (secondary).
4. Only patients with tumours of an invasive, primary, and malignant behavioural code (=3) in ICD 0 (http://www.who.int/classifications/icd/adaptations/oncology/en/) are included. It is worth noting that sometimes in revisions of ICD 0, the behavioural codes have been changed. In ovarian cancer some behavioural codes 3 were reclassified as 1 (uncertain behaviour) when moving from the 2nd to 3rd edition of ICD 0; this has relevance when comparing historic estimates to recent estimates.
5. Sex = Male or Female

5.2.2. Exclusion Criteria

1. Subsequent primary tumour registrations should be removed. As we are calculating person prevalence we are only concerned with the first tumour registered to a patient. If you are calculating tumour prevalence then this exclusion criteria should be ignored.
2. Exclude patient if the following information is missing or imputed: sex, date of diagnosis, date of birth or age.
3. Patients who are not resident in England at time of diagnosis. This criterion should minimise crossover and potential duplication of patients between different registries when looking at UK data rather than one specific nation.
4. Any patient who has died before the censor date (31st December 2015). As we are interesting in prevalence our cohort must be alive at the censor date to be included in the calculation. Any patient with a date of death after the censor date should be included in the analysis and their vital status should be set as alive for this purpose of this work.
5. Any patient with an embarkation date. Patients who have left the country at any point during the study should be removed from the analysis as there is no guarantee that we can ensure their vital status is correct at time of analysis.

5.3. Stratification Criteria

Whilst the categories that the data can be stratified by are the same variables or dimensions as for observed prevalence, there are slight differences in the details surrounding these categories. Details are listed below and are very similar to the methodologies used by Maddams et al. in 2009.[3]

- Geography
    - National Level
        - The modelling of data should be calculated on a national level to allow for variations in each nation to be accounted for.
        - UK level data can be calculated by combining the results of the individual countries.
    - Subnational breakdowns are not recommended due to the large numbers needed for modelling of early prevalence data.
- Age groups
    - Currently methodologies have only used age at diagnosis for complete cancer prevalence; this is due to difficulties in calculating the age at the end of follow up after already estimating age at diagnosis.
        - Broad age at diagnosis groups – 0-39, 40-69, 70+
            - Broader age groups allow for large enough cohorts for the modelling process. Potential to break down age groups where numbers are large e.g. 70+
- Tumour Group
    - Please see suggested tumour groupings in the appendix
- Sex

- 1 = males, 2 = females

Whilst granularity is lost using this method of stratification, it does allow an overall figure for people living with and beyond cancer to be determined for the UK.

5.4. Complete prevalence modelling methodology

When using a regression model to generate estimates of people diagnosed with cancer for each age group to cover the period before robust data was available (pre-1971 for England and Scotland, pre-1985 for Wales and pre-1993 for Northern Ireland), you need to take into account how far back the model could extend, while still meeting the inclusion and exclusion criteria detailed above.

For example, when estimating the number of people diagnosed with cancer aged 0-39 before 1971, a patient who is 39 years old at diagnosis (i.e. the oldest member of the age group) must not be older than 104 at the end of 2015. Therefore, the earliest year this patient could be diagnosed is 1950 (a patient who was 39 years old when diagnosed in 1950 would be 104 in 2015).

However, for the 40-69 age group, the oldest patient (69 years old at diagnosis) can be diagnosed no earlier than 1980 in order to be included (a patient diagnosed at 69 years old in 1979 would be 105 in 2015 and therefore too old and be excluded). This therefore prevents you from using modelled estimates for this group and any older age of group, such as a 70+ age group (where the year of diagnosis cut off would be 2010) for England. This is also true for Scotland (where data was available for the same period as England); however, for NI and Wales the 40-69 age group can only be modelled back to 1980 as data was only available from 1993 and 1985 respectively.

Similar to Maddams et al. (2009)[3], for the purposes of the regression model it is recommended that you use the following criteria:

Please note the following details are for England data, please change the below to align with the data availability from each nation.

- Include data on all patients diagnosed between 1 January 1971 and 31 December 2010.
  - Data on patients diagnosed between 2011 and 2015 (patients who will have survived for less than five years) should be excluded to avoid potential bias when estimating patients who had survived for over 40 years (patients diagnosed before 1971)
  - The excluded patients are used in the final count but are not included in the modelling process
- Data on prostate cancer cases are only used if they were diagnosed before 1 January 1992
  - This is to account for the effect that prostate specific antigen (PSA) testing had on the incidence rates of prostate cancer in England after it was introduced in 1992
  - This rule should also apply to Scotland; however, Wales and Northern Ireland should be approached differently. This is explained in the "Prostate Cancer 40-69 Age Group Modelling Approach" section – Section 5.5.

- A negative binomial regression model was used on each grouping of tumour site, sex and age group to estimate the number of cancers diagnosed before 1971 (same methodology used in Maddams et al. (2009)[3]).

    - Included a log link function
    - Dependent variable – "Prevalence Count on censor date"
    - Independent variable – "year since diagnosis" *
    - Offset – log ("country population (age group specific)")

*Please note in Maddams et al, (2009) the independent variables are listed as country and years since diagnosis. If you wish to do UK wide work then this can be applied, or the data can be split into data from each individual country and the UK figure can be achieved through adding the individual nation totals.

5.5. Prostate cancer – special considerations

When estimating complete prevalence for cancer we need to account for the effect of PSA testing on prostate cancer incidence and prevalence, which started around 1992 in England and Scotland. This approach was used in Maddams et al. (2009)[3]. Data for Wales and Northern Ireland were available starting from 1985 and 1993 respectively. As a consequence of this you can use a form of simple linear regression known as the 'annual percentage change' (APC) to determine the trend in prevalence of prostate cancer in men between the ages of 40-69 in the years 1971 to 1991 in England. The numbers for the 0-39 age group are normally deemed too small to model.

It should be assumed that the APC in the prevalence of prostate cancer in Wales and Northern Ireland are the same as England's and therefore this is applied to estimate the trends for both Wales and NI pre-1992 for pre-PSA patients.

The LINEST function in Excel that calculates the statistics of a line using the 'least squares' method was used to determine the APC. The APC for England is then applied to the earliest data point in both respective datasets (Wales and Northern Ireland), working back until the formula returns a prevalence estimate of below one patient or reached 1980 (the cut off for modelling the 40-69 age group as mentioned in the methodology section). This is where the data is then cut to ensure all criteria for the analysis were maintained.

The observed counts (including the years from 2011-2015 that were excluded from the regression) and the modelled estimates were collated and presented according to the various stratifications.

5.6. Data Disclosure Control Methods

Previous work has used the following criteria to adhere to national guidelines of data disclosure.

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.

As a general rule, suppression should occur when the underlying population of the area being used is below 1000 people however, please note that these are only guidelines and all suppression should be checked with each nation before publication. This also applies for final signoff of the release of this data.

If anything here is unclear or you feel that important information has not been included then we would like to hear from you. Please email: james.charnock@phe.gov.uk

Appendix

Oracle SQL script used for extracting prevalence data from English National Cancer Analysis System (CAS) can be provided on request.

Cancer groups used in phase two of the UK Cancer Prevalence Project

A - CANCER GROUPS

Cancer site/group	ICD10 codes
Breast	C50
Central Nervous System (including Brain)	C70, C71, C72, C751, C752, C753, D32, D33, D352, D353, D354, D42, D43, D443, D444, D445
Colorectal	C18, C19, C20
Endocrine	C73, C74, C75
Gynae (with Cervix In-situ)	C51, C52, C53, C54, C55, C56, C57, C58, D06
Gynae (without Cervix In-situ)	C51, C52, C53, C54, C55, C56, C57, C58
Haematology	C81, C82, C83, C84, C85, C88, C90, C91, C92, C93, C94, C95, C96
Head and Neck	C00, C01, C02, C03, C04, C05, C06, C07, C08, C09, C10, C11, C12, C13, C14, C30, C31, C32
Lower Gastrointestinal	C17, C18, C19, C20, C21, C26
Lung and Trachea	C33, C34
Malignant Melanoma	C43
Prostate	C61
Respiratory	C33, C34, C37, C38, C39, C45
Sarcoma	C40, C41, C46, C48, C49
Upper Gastrointestinal	C15, C16, C22, C23, C24, C25
Urology	C60, C61, C62, C63, C64, C65, C66, C67, C68, D090

B - MOST COMMON CANCERS

Cancer site/group	ICD10 codes
Bladder	C67
Breast	C50
Central Nervous System (including Brain)	C70, C71, C72, C751, C752, C753, D32, D33, D352, D353, D354, D42, D43, D443, D444, D445
Cervix	C53
Colorectal	C18, C19, C20
Head and Neck	C00, C01, C02, C03, C04, C05, C06, C07, C08, C09, C10, C11, C12, C13, C14, C30, C31, C32

Hodgkin Lymphoma	C81
Kidney	C64, C65, C66, C68
Leukaemia - Acute Myeloid	C920, C924, C925, C930, C940, C942
Leukaemia - Chronic Lymphocytic	C911
Liver	C22
Lung and trachea	C33, C34
Malignant Melanoma	C43
Multiple Myeloma	C90
Non-Hodgkin Lymphoma	C82, C83, C84, C85
Oesophagus	C15
Ovary	C56, C57
Pancreas	C25
Prostate	C61
Stomach	C16
Uterus	C54, C55

C - DETAILED SITE-SPECIFIC CANCERS

Cancer site/group	ICD10 codes
Anus	C21
Bladder (with in-situ)	C67, D090
Bladder (without in-situ)	C67
Breast	C50
Breast (with In-situ)	C50, D05
Cancer of Unknown Primary	C77, C78, C79, C80
Cervix	C53
Colorectal	C18, C19, C20
Gallbladder	C23
Head and neck - Eye	C69
Head and neck - Hypopharynx	C12, C13
Head and neck - Larynx	C32
Head and neck - Nasopharynx	C11
Head and Neck - Non-Specific	C00, C14, C31
Head and neck - Oral cavity	C02, C03, C04, C06
Head and neck - Oropharynx	C01, C09, C10
Head and neck - Palate	C05
Head and neck - Salivary Glands	C07, C08
Head and neck - Thyroid	C73
Heart, Mediastinum and Pleura	C38

Hodgkin Lymphoma	C81
Kidney and Unspecified Urinary Organs	C64, C65, C66, C68
Leukaemia - Acute Lymphoblastic	C910
Leukaemia - Acute Myeloid	C920, C924, C925, C930, C940, C942
Leukaemia - Chronic Lymphocytic	C911
Leukaemia - Chronic Myeloid	C921
Liver	C22
Lung	C33, C34
Melanoma	C43
Mesothelioma	C45
Multiple Myeloma	C90
Nasal Cavity and Middle Ear	C30
Non-Hodgkin Lymphoma	C82, C83, C84, C85
Oesophagus	C15
Ovary	C56, C57
Pancreas	C25
Penis	C60
Prostate	C61
Sarcoma - Connective and Soft Tissue	C49
Sarcoma - Retroperitoneum and Peritoneum	C48
Sarcoma: Bone	C40, C41
Small Intestine	C17
Stomach	C16
Testis	C62
Uterus	C54, C55
Vagina	C52
Vulva	C51

References

1. Support, M.C. *Cancer Statistics*. 2015; Available from: http://www.macmillan.org.uk/about-us/what-we-do/evidence/cancer-statistics.html.
2. *Cancer Prevalence UK Data Tables*. 2015: http://www.ncin.org.uk/view?rid=2954.
3. Maddams, J., et al., *Cancer prevalence in the United Kingdom: estimates for 2008.* Br J Cancer, 2009. **101**(3): p. 541-7.
4. Maddams, J., M. Utley, and H. Moller, *Projections of cancer prevalence in the United Kingdom, 2010-2040.* Br J Cancer, 2012. **107**(7): p. 1195-202.