Public Health England

In partnership with WE ARE MACMILLAN. CANCER SUPPORT

UK Complete Cancer Prevalence for 2013

Technical report

National Cancer Registration and Analysis Service and Macmillan Cancer Support in collaboration with the national cancer registries of Northern Ireland, Scotland and Wales

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. It does this through world-class science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. PHE is an operationally autonomous executive agency of the Department of Health.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Prepared by: James Charnock and Joanna Pethick
For queries relating to this document, please contact: ncrasenquiries@phe.gov.uk

Published September 2016
PHE publications gateway number: 2016267

Contents

Executive summary

The 2013 complete cancer prevalence project presents an update of the estimates published by Maddams et al in 2009[1], relating to the number of patients who were alive at the end of 2008 who had previously been diagnosed with cancer. Cancer prevalence is a product of both cancer incidence and survival trends; it is a useful statistic to help us understand the longer term effects of cancer and its treatment.

This new work on cancer prevalence forms part of the Macmillan Cancer Support and PHE's National Cancer Registration and Analysis Service (NCRAS) partnership[2]. It builds on previous publications from this partnership, involving the analysis of 20-year prevalence of cancer in the UK at the end of 2010.

This project was developed in collaboration with the population-based cancer registries for Northern Ireland, Scotland and Wales. We have now updated the number of patients living after a cancer diagnosis in the UK, up to the end of 2013. This timeframe was aligned with the availability of cancer registration data across the UK nations.

To estimate the number of patients diagnosed with cancer many years ago and prior to the establishment of cancer registries in each nation, we constructed a negative binomial regression model.

The estimated total number of patients living with a cancer diagnosis at the end of 2013 was 2.27 million in the UK: 82% were from England and 9% from Scotland, while Wales and Northern Ireland contributed 6% and 3% respectively to the total.

The workbook containing a breakdown of the numbers of people alive at the end of 2013, stratified by cancer site (of first diagnosis of cancer if patients have more than one cancer diagnosis), year of diagnosis, age, sex and time since diagnosis, has been published and can be found at www.ncin.org.uk/about_ncin/segmentation.

Objectives

Maddams et al 2009[1] estimated that there were 2 million people living with cancer in the UK at the end of 2008. This piece of work was extended to project future prevalence figures, in Maddams et al 2012[3]. Additional analysis by Macmillan Cancer Support based on Maddams et al 2012 estimated that there would be 2.5 million people living with a cancer diagnosis in the UK in 2015.

As part of the Macmillan-NCRAS UK Cancer Prevalence Project, phases 1 and 2 looked at limited duration 20-year prevalence up to 2010. We have now extended this initial work to incorporate more recent data and to estimate complete prevalence (the total number of people living with cancer in a population at a given point in time) in the UK at the end of 2013 (with an index date of 31 December 2013).

This data will also be used to inform a programme of work led by Macmillan Cancer Support that aims to project cancer prevalence up to 2040.

Methods

Cancer registry data was available from all the UK nations, however the data related to different periods of diagnosis for each nation.

Data was available covering the following period of diagnosis for the respective nations from their population based cancer registries:

- England: 1971 to 2013
- Scotland: 1971 to 2013
- Wales: 1985 to 2013
- Northern Ireland (NI): 1993 to 2013

The methodology for calculating the complete prevalence (defined as the sum of observed prevalence, from the years of diagnosis that were available in the registry data, in addition to modelled prevalence, from those that were not), described below, focuses on England data. However, the methodology is largely the same across all nations. Any differences between methodologies are described in the section below.

Observed data methodology

Data for all patients diagnosed with cancer between 1 January 1971 and 31 December 2013 was extracted from the PHE Cancer Analysis System, for England.

The following inclusion and exclusion criteria were applied to identify the records of patients who were still alive on 31 December 2013:

1. Only include patients resident in England at the time of diagnosis
2. Exclude any patient with an “Embarkation Date”
3. Exclude patients who were older than 99 years of age at diagnosis
4. Exclude patients who were older than 104 years of age at the censor date (31 December 2013)
5. Exclude any patients who died prior to and including 31 December 2013
6. Only the patient’s first tumour diagnosis was included in the extraction
    - this counts number of people with cancer rather than the number of tumours, so figures herein relate to person prevalence rather than tumour prevalence

The data was then classified using the following criteria:

- tumour site of the first tumour diagnosis – the four most common cancers by incidence: Breast ICD-10 C50 (female only); Prostate C61 (male only); Colorectal C18-C21; and Lung C33-C34; in addition to “All Others” group (C00-C97 excluding C44)
- the “All Others” group includes male breast cancer and only includes patients not in any of the four most common cancers.
- sex (1= males, 2= females)
- age group at diagnosis (0-39, 40-69 and 70+)
- year of diagnosis (eg 2010)
- actual prevalence – number of patients for each of the above groupings

Criteria 1, 2, 3 and 4 were used to extract and enumerate the number of cancers diagnosed during the same period including those who have died since diagnosis.

Modelling methodology

The data was stratified in order to stay as close as possible to Maddams’ original methodology. The statistical modelling involved estimating the number of people diagnosed with cancer before 1971, which covers the time period before the official establishment of the English national cancer registration system.

When generating the regression model for each age group to cover the period pre-1971, we had 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 (ie the oldest member of the age group) must not be older than 104 at the end of 2013. Therefore, the earliest year this patient could be diagnosed is 1948 (a patient who was 39 years old when diagnosed in 1948 would be 104 in 2013).

However, for the 40-69 age group, the oldest patient (69 years old at diagnosis) can be diagnosed no earlier than 1978 in order to be included (a patient diagnosed at 69 years old in 1977 would be 105 in 2013 and therefore too old and be excluded). This therefore prevented us from using modelling estimates for this group and also the 70+ age group (where the year of diagnosis cut off is 2008) for England, which was covered by the national cancer registration system from 1971.

- 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 be modelled back to 1978 as data was only available from 1993 and 1985 respectively

Similar to Maddams et al 2009[1], for the purposes of the regression model we have used the following criteria:

- include data on all patients diagnosed between 1 January 1971 and 31 December 2008
- data on patients diagnosed between 2009 and 2013 (patients who will have survived for less than five years) were removed 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
- prostate cancer cases were only used if they were diagnosed before 1 January 1992
- this was done 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 was also applied to Scotland; however, Wales and Northern Ireland were approached differently. This is explained in the “Prostate Cancer 40-69 Age Group Modelling Approach” section
- 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)
- included a log link function
- dependent variable – “Prevalence Count”

- independent variable – “Diagnosis Year”
- offset – log (“Incidence Count”)

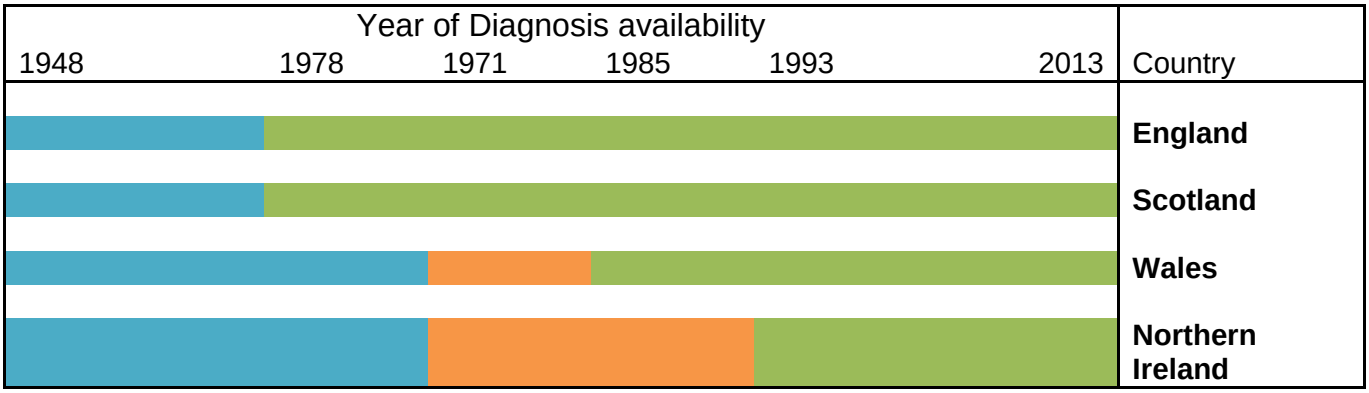
An assumption that the incidence count in 1971 for each tumour site and sex grouping in the 0-39 age group was the same in all preceding years back to 1948 was used to generate the offset.

The model outputs and comparison between the observed data and modelled data are shown in the appendix.

Figure 1 shows the source of the data for each country by the following three categories:

Green – Observed data for all age groups
Orange – Modelled data for 0-39 and 40-69 age groups (no data available for 70+ group)
Blue – Modelled data for 0-39 age group (no data available for 40-69 and 70+ age group)

Figure 1. Source of complete prevalence data availability broken down by country



Prostate cancer 40-69 age group modelling approach

As part of this project we are assuming that the effects of PSA testing on prostate cancer incidence and prevalence started in 1992. This is in line with Maddams et al (2009)[1]. Data availability for Scotland was the same as England; however, as data for Wales and Northern Ireland starts from 1985 and 1993 respectively, we have very little data available for these nations. As a consequence of this we used 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 were deemed too small to model the 0-39 age group.

It is 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 used to estimate the trends for both Wales and NI pre-1992 too.

The LINEST function in Excel that calculates the statistics of a line using the 'least squares' method was used to determine the APC.

The observed data in Figure 2 below refers to the counts of patients who were diagnosed in each year with prostate cancer in England between the ages of 40-69 at diagnosis between 1971 and 1991 and were still alive at the end of 2013 (31 December 2013), ie prevalence counts.

Figure 2. Comparing LINEST prevalence estimates to observed prevalence data for prostate 40-69 age group in England

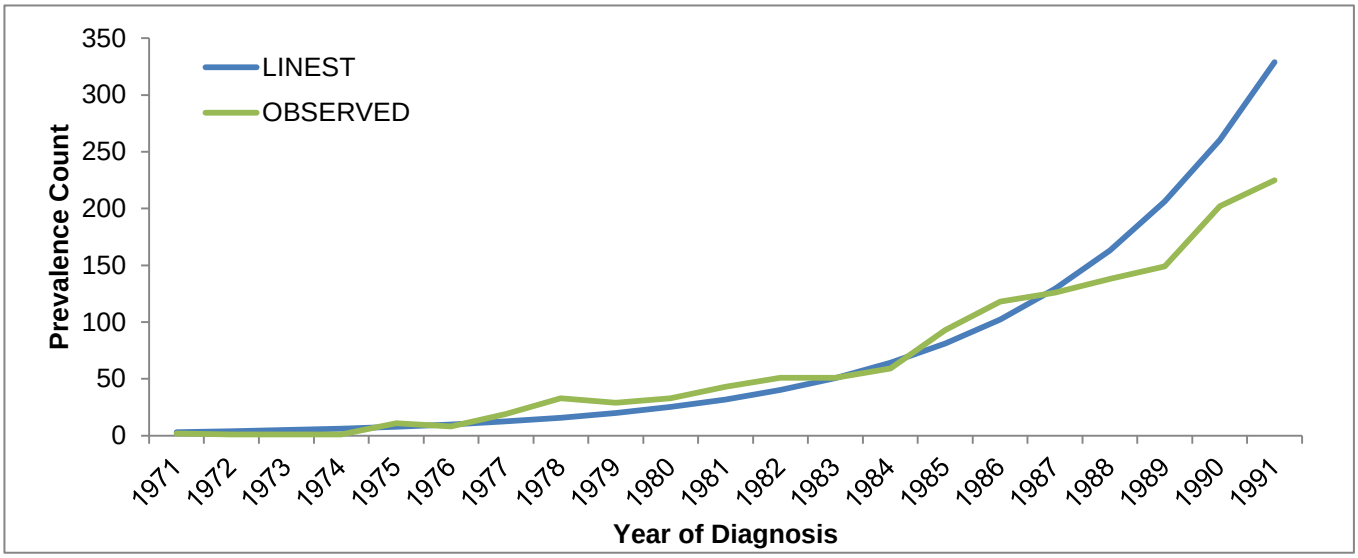
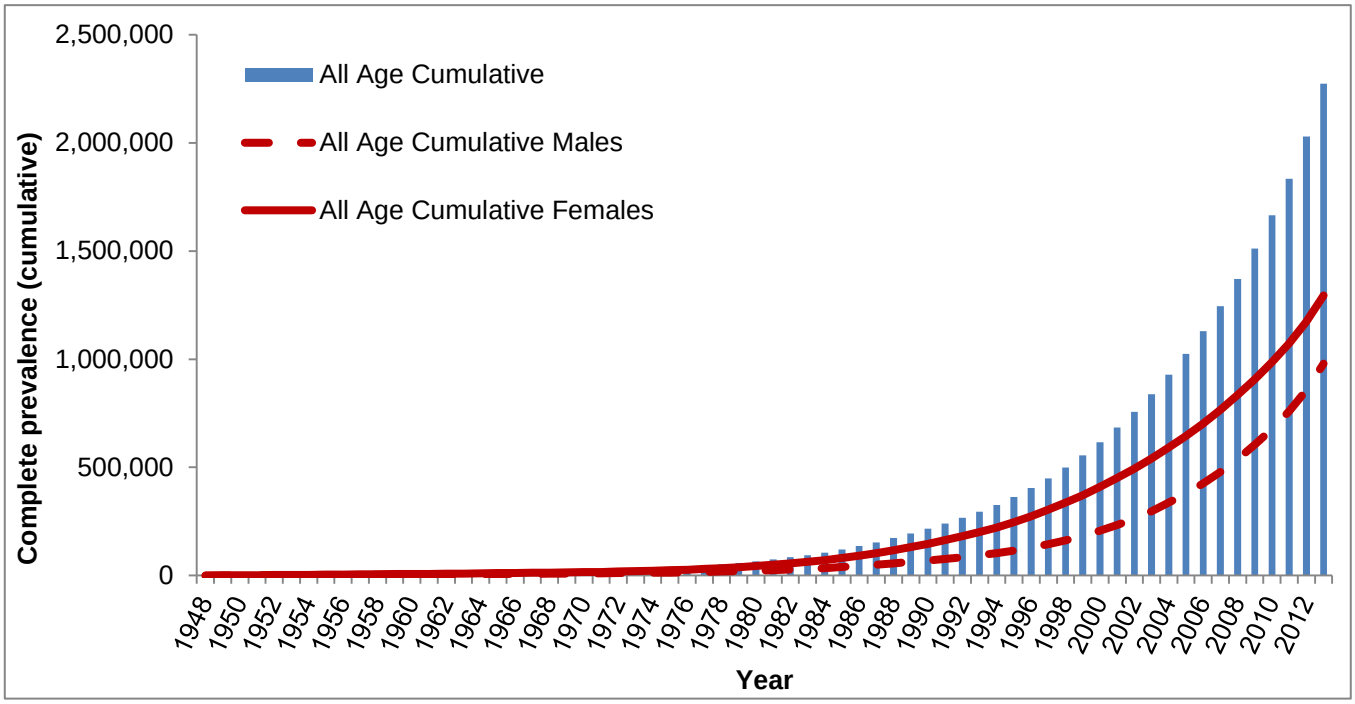


Figure 2 shows the predicted values calculated using the LINEST method compared to the observed data. The APC for England was then applied to the final data point in both respective datasets (Wales and Northern Ireland), working back until the formula returns a prevalence estimate of below one patient or reached 1978 (the cut off for modelling the 40-69 age group as mentioned in the methodology section). This is where the data was then cut to ensure all criteria for the analysis were maintained.

Results

We calculated that there were ~2.27 million people living with cancer in the UK at the end of 2013 (57% women and 43% men) (Figure 3). This is in line with estimates predicted by Maddams et al 2012 [3].

Figure 3. Complete cancer prevalence, the cumulative numbers of people living with cancer and who were diagnosed between 1948 and 2013, by year of diagnosis and sex, in the UK at end of 2013



England has the highest number of patients living with cancer, owing to its relatively large population size. Similarly, Northern Ireland has the lowest number, probably due to it having the smallest population size. When adjusted to account for these differing population sizes, the nation with the highest crude prevalence rate per 100,000 people is Wales, with England the lowest. This differs slightly from the work carried out by Maddams et al 2009[1], which showed that Northern Ireland had the lowest crude rates of all four nations. This does not account for differing age distributions across nations, which could reveal that the older population in Wales is a contributing factor to this trend.

Table 1 presents complete prevalence estimate figures at the end of 2013 broken down by cancer site, focusing on the four most common cancers by incidence (female breast, prostate, lung and colorectal). Lung was the smallest group (3% of total complete prevalence) reflecting poor prognosis associated with this cancer type. Female breast cancer was the most prevalent (26%) across the whole of the UK, owing partly to high

incidence as a result of earlier diagnosis from screening programmes combined with higher cancer survival as a result of effective breast cancer treatments. The number of years that have passed since diagnosis varies across each of the four cancer sites; however the overall trend shows that the majority of people living with cancer were diagnosed between 2009-2012 (1-5 years prior to the end of 2013).

The majority of people living with cancer – around 1,400,000 people – were diagnosed between the ages of 40 and 69, accounting for more than three-fifths of the complete prevalence cohort (61%). There were 310,000 people diagnosed in the younger age group aged 0-39, equating to 14% of all cancer survivors. One quarter of all people living with cancer, 565,000 patients, were diagnosed at age 70 and over.

The lower proportion of cancer survivors diagnosed in the 0-39 age group may be indicative of the fact that two thirds of cancer diagnoses occur in patients over 65 years of age[4, 5].

It may be possible that the comparatively lower proportion of cancer survivors in the 70+ age group is correlated with the relatively small number of people over 70 in the population (only 12% of the UK population are 75 and over)[6] poorer survival, later presentation at diagnosis and issues related to old age such as comorbidities and low performance status; but we cannot conclude that from this data release alone.

Table 1. Breakdown of UK complete prevalence at the end of 2013 by time since diagnosis

	Time since diagnosis (years)					
Cancer site	**0-1**	**1-5**	**5-10**	**10-20**	**>20**	**Total**
All malignant neoplasms (excluding non-melanoma skin cancer)	243,700 (11%)	785,200 (35%)	487,300 (21%)	490,000 (22%)	266,900 (12%)	2,273,200 (100%)
Female breast	45,500 (8%)	184,000 (31%)	137,000 (23%)	158,000 (27%)	69,900 (12%)	594,500 (100%)
Prostate	41,500 (12%)	152,300 (46%)	90,800 (27%)	46,300 (14%)	2,800 (1%)	333,800 (100%)
Colorectal	29,800 (11%)	102,700 (38%)	57,200 (21%)	56,600 (21%)	22,300 (8%)	268,600 (100%)
Lung	20,000 (31%)	25,000 (39%)	7,500 (12%)	6,500 (10%)	5,100 (8%)	64,200 (100%)
All other malignant neoplasms (excluding non-melanoma skin cancer)	106,800 (11%)	321,200 (32%)	194,800 (19%)	222,500 (22%)	166,800 (16%)	1,012,100 (100%)

Based on first cancer diagnosis only, rather than the first specific cancer diagnosis method used in phase 2 of the UK Cancer Prevalence Project.

Limitations

Currently these estimates are presented at a national level with no breakdowns at a sub-national geography. Age groups are also limited to the three broad age groups to enable a robust modelling approach.

Finally, the data presented here is only broken down by the four most common cancers in terms of incidence (female breast, prostate, colorectal and lung). Future work at a more detailed level (expanding to a larger number of tumour sites) would enable a better understanding of the cancer population.

Conclusions

The provisional cancer prevalence figure for the end of 2013 was as estimated by Maddams et al 2012[3]. Our analysis provides a more up-to-date understanding of the UK population of people living with cancer using the latest available data.

This will feed into a further work programme led by Macmillan Cancer Support that aims to project the cancer prevalence population up to 2040. This will allow an understanding of the potential future trends associated with people living with cancer in the UK and is important for planning for future distributions of resources and services for these people.

Acknowledgments

This technical document and workbook were prepared by James Charnock and Joanna Pethick, who are members of the Macmillan Cancer Support/NCRAS partnership, and Samuel Jones who works for Macmillan Cancer Support.

This work is part of the Macmillan Cancer Support and PHE's NCRAS Partnership Work Plan. The data sourced is presented in collaboration with: the Welsh Cancer Intelligence and Surveillance Unit; the Health Intelligence Division, Public Health Wales; the Scottish Cancer Registry; and the Northern Ireland Cancer Registry, which is funded by the Public Health Agency, Northern Ireland.

NCRAS wishes to thank all those who have contributed to the content and production of this publication, and without whom the work would not have been possible.

Special thanks must also go to the following contributors who had key inputs into the project:
Professor Henrik Møller – King's College London
Sarah Miller – PHE
Dr Katherine Henson – PHE
Dr Isabella Carneiro – PHE/Cancer Research UK
Vivian Mak – PHE

References

1. Maddams, J., et al., *Cancer prevalence in the United Kingdom: estimates for 2008.* Br J Cancer, 2009. **101**(3): p. 541-7.
2. (NCRAS), M.C.S.N.C.R.a.A.S. *NCRAS/Macmillan Partnership*. Available from: http://www.macmillan.org.uk/about-us/what-we-do/evidence/evidence-department/our-partnerships/national-cancer-registration-and-analysis-service.html#271857.
3. 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.
4. NCIN, N.C.I.N.-. *Older people and cancer*. 2015; Available from: http://www.ncin.org.uk/publications/older_people_and_cancer.
5. Statistics, O.f.N. *Cancer Registration Statistics, England Statistical bulletins*. 2014; Available from: http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/previousReleases.
6. Statistics, O.f.N. *Population Estimates Summary for the UK, mid-2015*. 2016; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland.

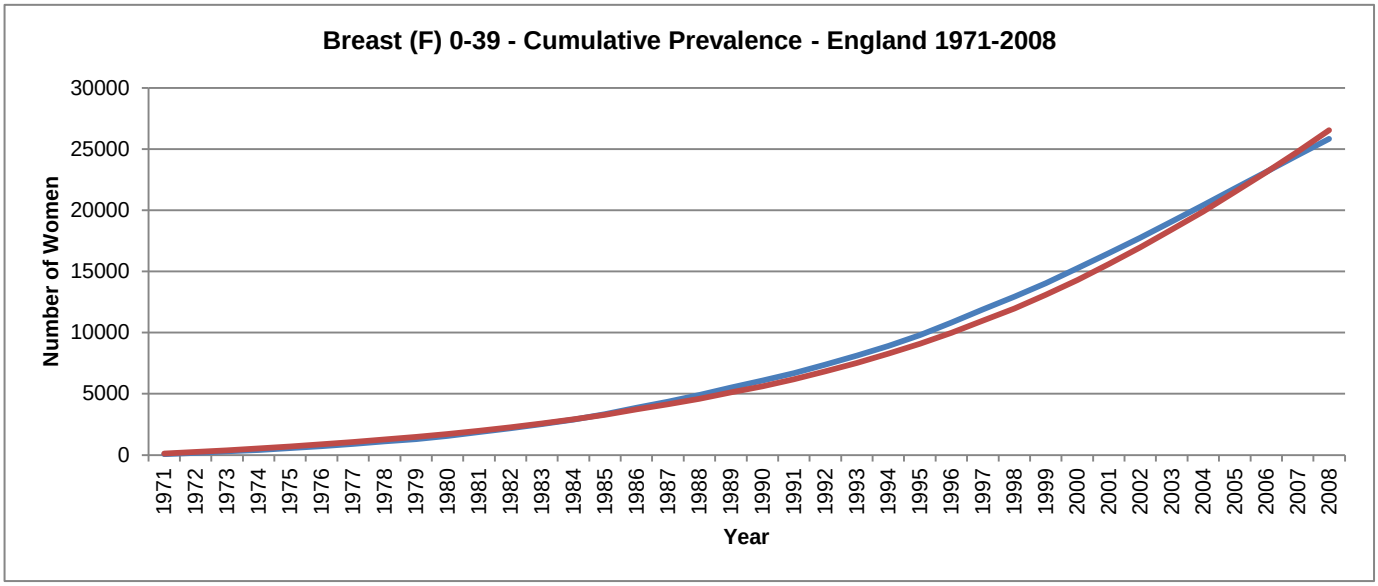
Appendix

Negative binomial regression outputs by tumour site for England

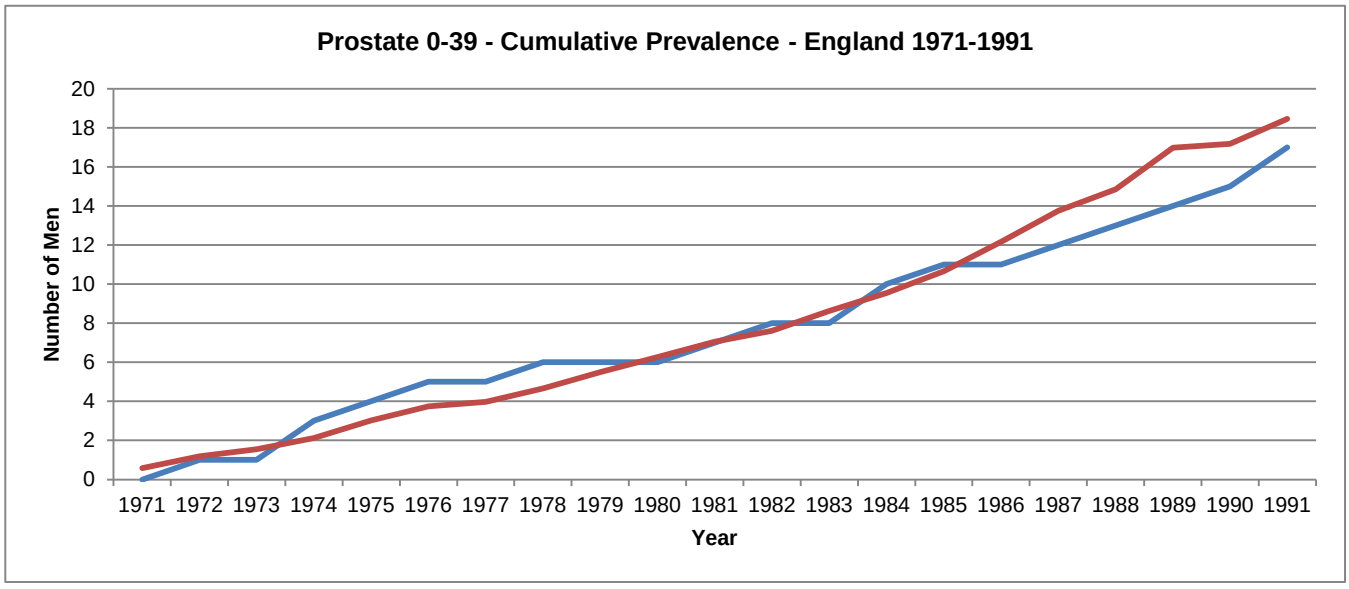
The graphs represent the observed counts (**in blue**) and the modelled counts (**in red**). The graphs show a time period of 1971 to 2008 (except for Prostate (M) 1971 to 1991), as this was the best way to compare the outputs from the model with the observed counts. For the final calculation, data back to 1948 for the 0-39 age group was calculated.

These graphs show a cumulative count of the prevalence of patients diagnosed who were still alive on 31 December 2013. This means that the data point for 1980 includes all the patients diagnosed in 1980 and before, who were still alive at the censor date.

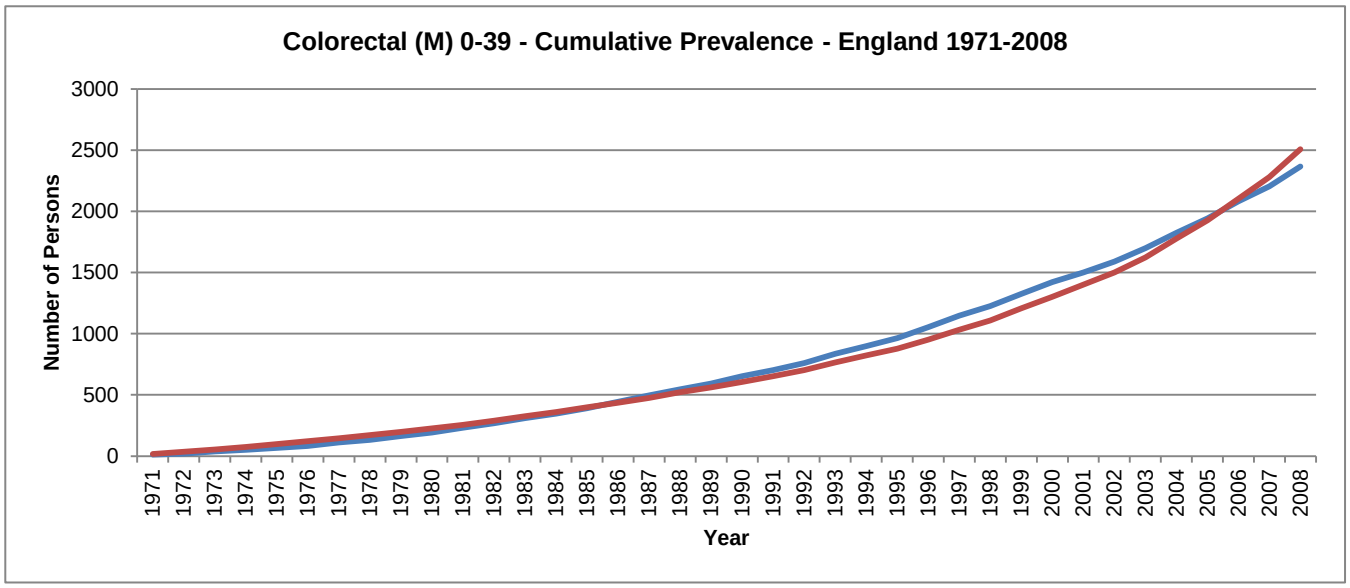
Female Breast Cancer



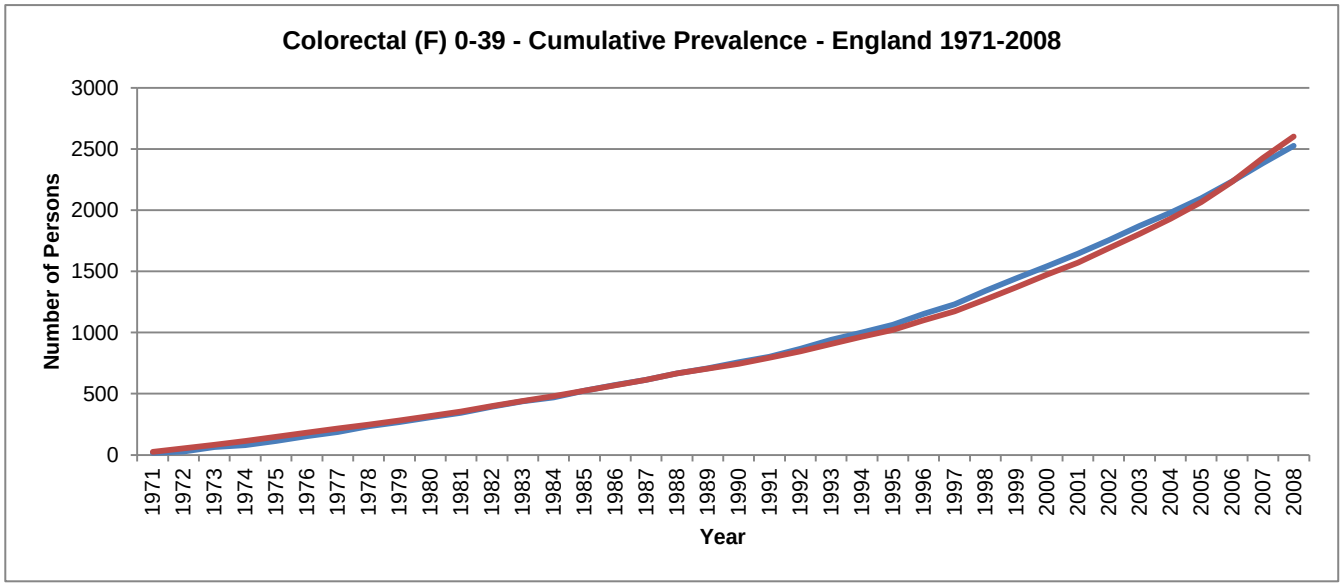
Prostate Cancer



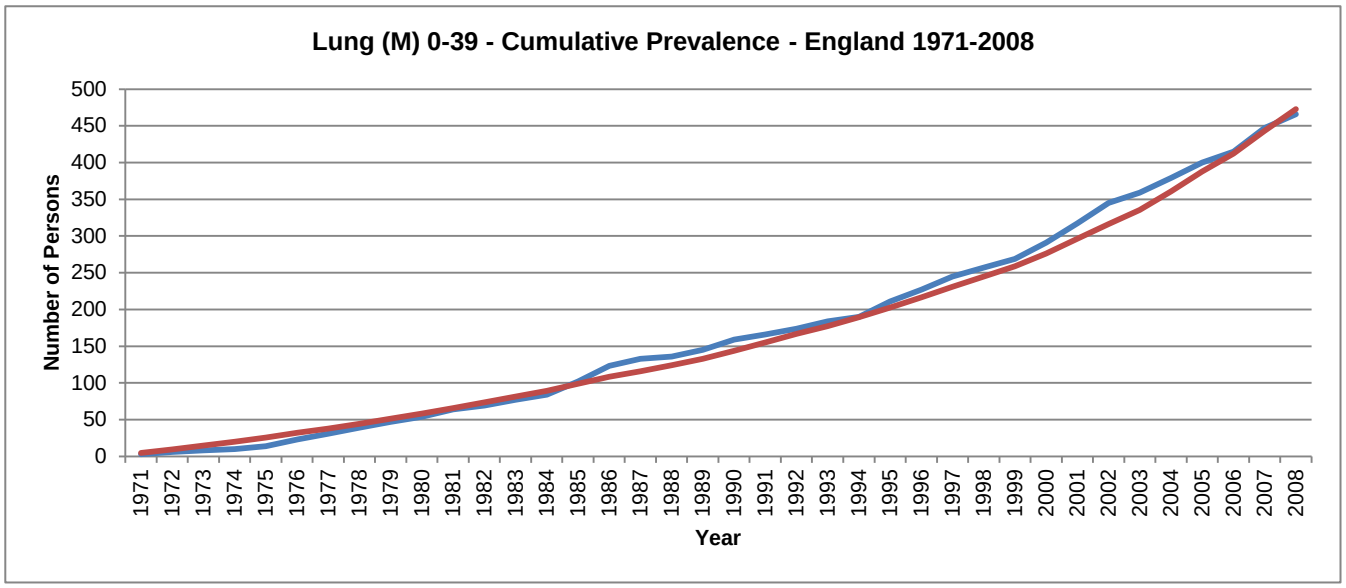
Male Colorectal Cancer



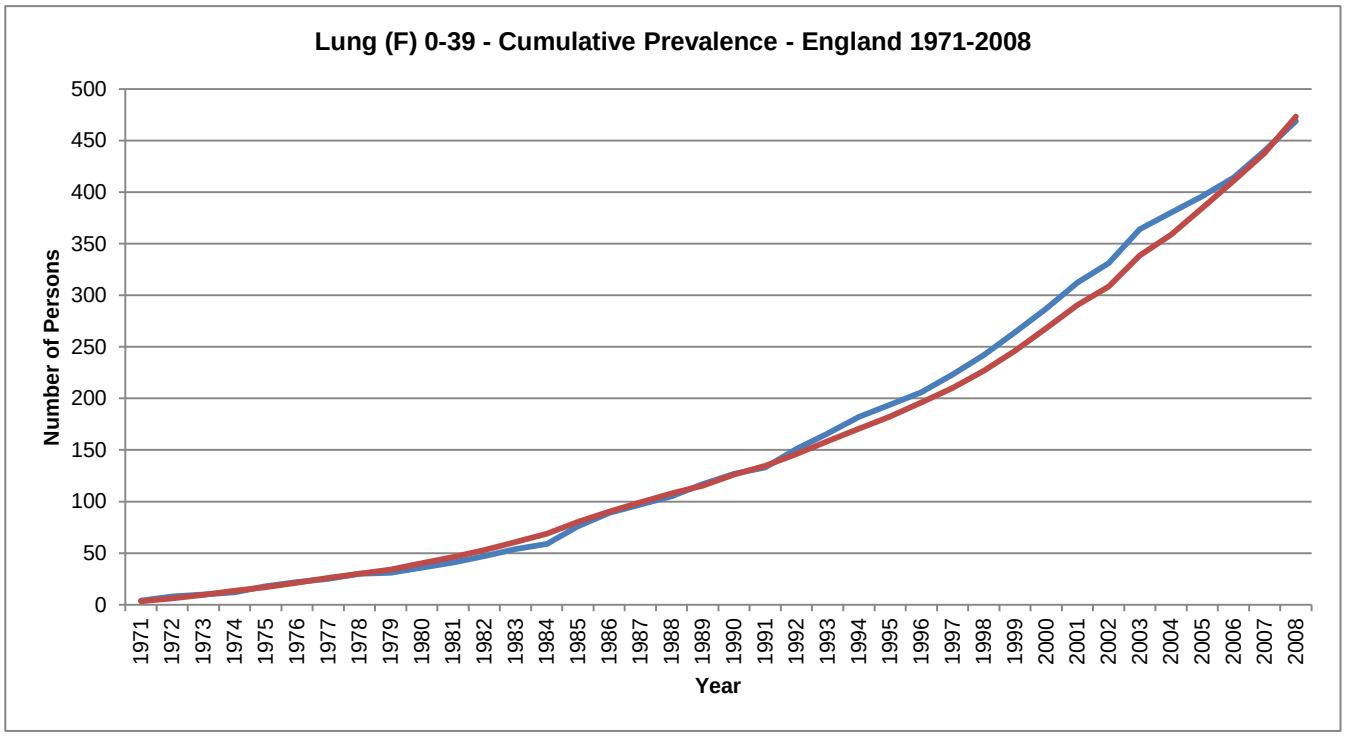
Female Colorectal Cancer



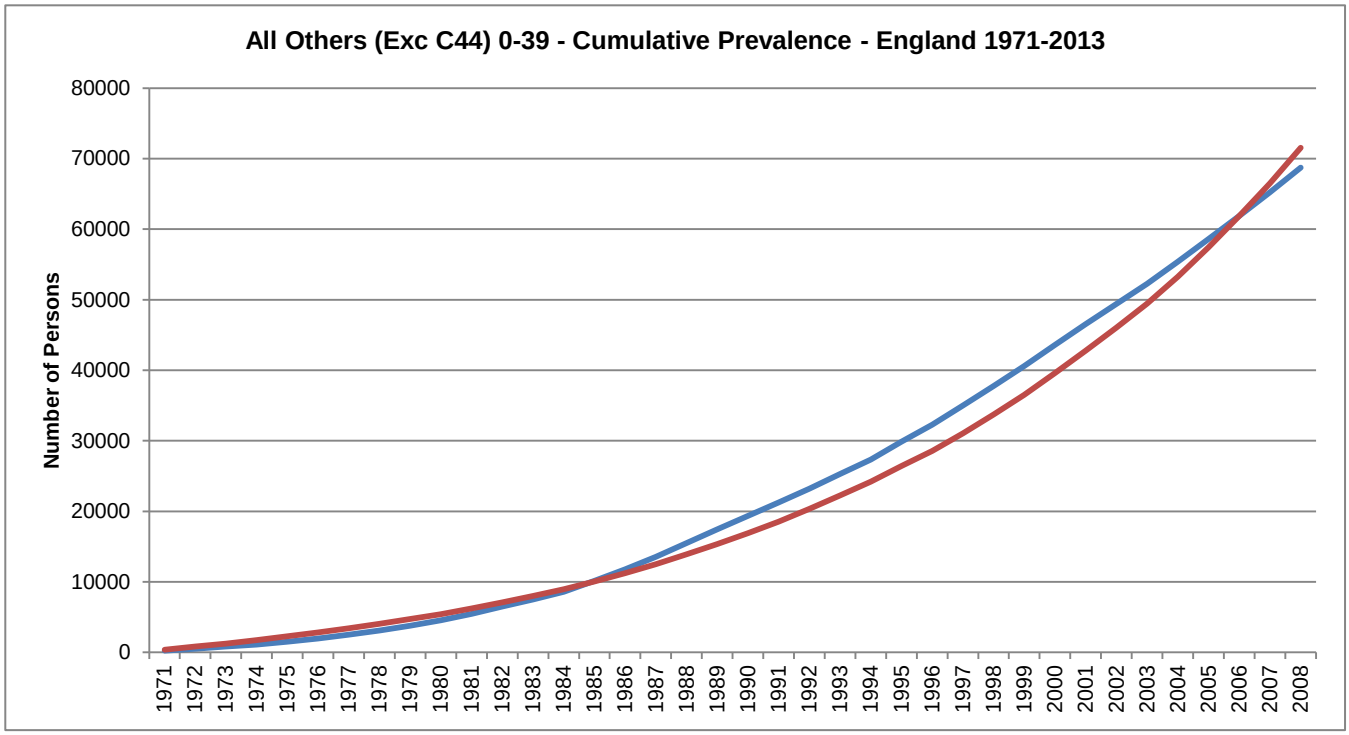
Male Lung Cancer



Female Lung Cancer



Male All Others (Exc C44)



Female All Others (Exc C44)