Get Data Out cancer survival analysis technical document, November 2019

This analysis follows the UK and Ireland Association of Cancer Registries (UKIACR) ratified Standard Operating Procedure Guidelines on Population Based Cancer Survival Analysis; these are the same guidelines that are used to produce the National Statistics on cancer survival in England. The patient information used for the survival analysis was extracted from AV2017.AV_TUMOUR on CASREF01. The Cancer Survival SOP v11_0 was followed.

Net survival estimates the survival of cancer patients compared with the background mortality that patients would have experienced if they had not been diagnosed with cancer. Net survival is a variant of relative survival that is preferred as a measure of cancer survival in adults because it is an unbiased estimator.

The survival estimates have been calculated using the Pohar-Perme net survival estimator, as implemented by stns in Stata 15. The Pohar-Perme estimator of net survival is an unbiased estimator that accounts for informative censoring bias. Background mortality is derived from population life tables. The life tables used are supplied by the Office for National Statistics. When using these lifetables, the mortality of cancer patients is compared with that of individuals in the general population who belong to the same single year of age (0 to 99 years), sex, population weighted quintile of the index of multiple deprivation (IMD) and region.

Further information on the quality and methodology information (QMI) can be found in the Cancer survival statistical bulletins QMI.

In the survival analyses, for age groups where the estimates do not meet the following quality criteria, the results are suppressed for those particular age groups of the specific cancer site:

A minimum of ten patients should be alive at the beginning of the survival period being estimated (for example, first year of follow-up for a 1-year estimate; no cohorts failed this criterion)

At least two deaths registered in the years before or after the duration(s) being estimated (cohorts failing this criterion are denoted ".f" in the results)

The standard error of the survival estimates should be lower than 20% (cohorts failing this criterion are denoted ".g" in the results)

The level of the survival estimates should not increase with duration; for example, the survival estimated at 6-months following diagnosis should be lower than the survival estimated at 3-months following diagnosis (cohorts failing this criterion are denoted ".h" in the results)

Survival analysis data quality criteria were implemented to identify the patients that were eligible to be included in the analysis. Other decisions applied include:

- 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
- when two or more tumours of the same type are diagnosed on the same day for a patient, the one with the worst prognosis is chosen for inclusion

Cancer in adults is defined using the International Statistical Classification of Diseases 10th Revision (ICD-10) and by morphology and behaviour codes in the International Classification of Diseases for Oncology, Second Edition (ICD-O-2). Get Data Out metrics include some cancers with ICD-10 codes that in ICD-10 are not classified as potentially lethal. These cancers have been excluded from survival analysis and include D39.1 for ovarian cancers and D29 for testicular cancers. This exclusion is in line with other survival analyses produced by the National Cancer Registration and Analysis Service (NCRAS).

Changes from previous release

In previous Get Data Out (GDO) releases, only data from the GDO cohort (patients diagnosed in 2013-2016) was used for survival analysis. This did not account for diagnoses in these patients that occurred prior to 2013, and thus in some instances secondary tumours were counted as primary tumour diagnoses. This year, historic data was included in the analysis allowing all secondary tumour diagnoses to be excluded, giving a more accurate estimation of survival.

In the previous release (GDO_0011, 2019-02-18) there was an error in the calculation of crude survival confidence intervals for brain, meningeal and other primary CNS tumours. This has been corrected in the current release.