

Background notes for net survival and probability of cancer death

1. Net survival is an estimate of the probability of survival from the cancer alone. It can be interpreted as the survival of cancer patients after taking into account the 'background' mortality that the patients would have experienced if they had not had cancer. Background mortality is derived from life tables of all-cause mortality rates in the general population. Net survival varies with age, and the age profile of cancer patients can vary with time and between geographical areas, so the estimates are also age-standardised to facilitate comparison. Estimates are shown with their 95% confidence intervals (see Background Note 7). For convenience, net survival is expressed as a percentage in the range 0–100%.
2. The period approach to estimate survival is used to predict the survival for patients who were diagnosed more recently, and for whom long-term follow-up data are not yet available. This approach is analogous to the estimation of life expectancy at birth for babies born last year. Those babies have not yet had an opportunity to be followed up for a lifetime to determine their "true" (cohort) life expectancy, which will only be known after a century. To calculate life expectancy at birth, the most recently available mortality rates for each single year of age up to 99 years are used. The period estimate of survival uses the same approach under the same assumptions. A period estimate of 5-year survival is a short-term prediction of survival for patients diagnosed in that period, assuming that they will experience the most recently observed conditional probabilities of survival in each year up to 5 years since diagnosis. The period approach results presented here are for all patients with any follow-up in the period 1st January 2009 to 31st December 2013. In order to compute these predictions, data are required for all patients diagnosed during 1999–2013.
3. We use an unbiased estimator of net survival (Pohar Perme et al. 2012). Patients known to have died on the same day as they were diagnosed (zero follow-up time) are included in the analyses. Patients for whom a death certificate was the only information available are excluded, because their duration of survival is unknown.
4. The probabilities of cancer death and non-cancer death were estimated for adults in England diagnosed during 2009–2013 using the period approach. Probability of cancer death is calculated by first estimating the amount of hazard experienced by cancer patients that is due to their cancer. This is done by subtracting the hazard which is expected due to population causes from the observed total hazard experienced by the patients, leaving the "excess" or cancer hazard. The probability of cancer death at a time t is then estimated as the overall probability of survival to time t multiplied by the cancer hazard at time t . This is because to experience cancer death at time t the patient must survive until time t and then die due to the cancer hazard. The cumulative probability that a patient dies from cancer by time t is then calculated as the sum of probabilities of cancer death in each interval of follow-up from the first day of follow-up to time t . The probability of non-cancer death is calculated in exactly the same way, except with the background population hazard used instead of the cancer hazard.
5. All adults (aged 15–99 years) in England who were diagnosed during 1999–2013 with one of 25 common or 13 less common cancers as an invasive, primary, malignant neoplasm were eligible for analysis. Patients whose tumour was benign (not malignant) or in situ (malignant but not invasive) or of uncertain behaviour (uncertain whether benign or malignant), or for which the organ of origin was unknown, are excluded. Details of the eligibility and exclusion criteria have been published (Li R., Abela L., Moore J., et al. 2014).
6. Cancers were defined by their anatomical location (site), morphology and behaviour (benign, in situ, or invasive). Tumour site was coded according to the 10th revision of the International Classification of Diseases (ICD-10). Morphology and behaviour were coded according to the second edition of the International Classification of Diseases for Oncology (ICD-O-2). Standard criteria were used to decide whether a tumour record was eligible for inclusion in the analyses (Li R., Abela L., Moore J., et al. 2014). Records were excluded if they contained data of inadequate quality or were for patients not resident in England.
7. All net survival estimates presented are accompanied with the respective 95% confidence intervals with the abbreviation "CI". A 95% confidence interval is an interval around an estimate (in this case a net survival estimate) which has a 95% chance of containing the true population value that it is estimating. It can be thought of as the range of values the true population value is very likely to lie in.
8. When the data for this report were extracted for analysis on 17 May 2015, cancer registrations up to 2013 were believed to be at least 99% complete, and each patient's vital status at 31 December 2014 was known for 99.5% of cancers registered during the period 2009–2013.

9. Issues of small numbers arise when examining long-term survival and rarer cancers. The following rules were applied to exclude estimates that are very unlikely to be insightful for patients:

- a) Estimates are not reported where data are too sparse. Specifically, survival at time T is only reported if there are 10 patients alive at T and if there were at least 5 events in either the period before or the period after T , and additionally at least one event in the 2 years before T . Estimates are therefore excluded where the patient numbers are too small to be robust.
- b) Net survival estimates are not reported where the range of the confidence interval is greater than 20% (e.g. 95% CI: 50%-71%). Large uncertainty for estimates makes results difficult to interpret for patients, and gives little insight, particularly as the estimated difference between 1-year and 10-year survival itself is frequently less than 20%.
- c) 1- and 5-year survival estimates are reported for patients aged up to 90, while 10-year survival estimates are restricted to patients age up to 70. No net survival estimates have been calculated for patients over 90. It is inappropriate to estimate long-term net survival of older patients due to the very large weight one or two surviving patients can have on the results, particularly when the initial cohort is small.
- d) If a net survival estimate is not reported due to one of the exclusion rules a), b) or c) then the corresponding probability of cancer death and probability of non-cancer death estimates are also excluded.

References

- Li R., Abela L., Moore J., et al. (2014) "Control of data quality for population-based cancer survival analysis." *Cancer Epidemiol.* 38, 314-320, doi:10.1016/j.canep.2014.02.013 (2014).
- Pohar Perme M., J. Stare and J. Estève (2012). "On estimation in relative survival." *Biometrics* 68: 113-120.

ICD-10 codes

Rarer Cancers		More Common Cancers	
Cancer	ICD-10 codes	Cancer	ICD-10 codes
Tongue	C01, C02 (not C02.4)	Oesophagus	C15
Oral Cavity	C03, C04, C05, C06	Stomach	C16
Salivary Glands	C07, C08	Colon	C18
Oropharynx	C09, C10, C02.4	Colorectum	C18-C20, C21.8
Hypopharynx	C12, C13	Rectum	C19 - C20, C21.8
Small bowel	C17	Liver	C22
Anus	C21 (not C21.8)	Pancreas	C25
Gallbladder	C23, C24	Larynx	C32
Nasal sinuses	C30, C31	Lung	C33, C34
Bones	C40, C41	Melanoma of skin	C43
Vagina	C51, C52	Mesothelioma	C45
Penis	C60	Breast	C50
Eye	C69	Cervix	C53
		Uterus	C54, C55
		Ovary	C56, C57.0 - C57.7
		Prostate	C61
		Testis	C62
		Kidney	C64 - C66, C68
		Bladder	C67
		Brain	C71
		Thyroid	C73
		Hodgkin lymphoma	C81
		Non-Hodgkin lymphoma	C82 - C85
		Myeloma	C90
		Leukaemia	C91 - C95

Sources: World Health Organization. *International Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10)* and *International Classification of Diseases for Oncology, Second Edition (ICD-O)*. Geneva: World Health Organization.