

Glossary

Age-specific rate: The rate that events occur per 100,000 persons of a particular age class. It is calculated using the formula:

$$a_i = \frac{r_i}{n_i} \times 100,000$$

where a_i is the age-specific rate for age class i , r_i is the number of events in age class i and n_i is the number of person years of observation in age class i .

Age-standardised rate (ASR): The rate per 100,000 persons that has been adjusted to take account of different age structures between geographic areas or time periods by adopting a reference population. It is calculated by the direct method using the following formula:

$$ASR = \frac{\sum_{i=1}^A a_i w_i}{\sum_{i=1}^A w_i}$$

where ASR is the age-standardised rate, a_i is the age-specific rate for age class i , w_i is the standard population of age class i and A is the number of age intervals.

Annual percentage change (APC): The percentage increase or decrease per year in the age-standardised rate (ASR). It is calculated by fitting a regression line to the natural logarithm of the rates using calendar year as a regression variable, i.e. $y = mx + b$, where $y = \ln(ASR)$, x is the calendar year, and b is a constant. The annual percentage change (APC) is thus given by

$$APC = 100 \times (e^m - 1)$$

The calculation assumes that the age-standardised rates increase or decrease at a constant rate over the period examined.

Behaviour: The manner in which a tumour acts, i.e. benign, in situ or malignant.

Benign tumour: A tumour that neither invades nor destroys the tissue in which it originates, nor spreads to distant sites in the body.

Cancer: A disease resulting from the breakdown in the normal growth of body cells as a result of faults or damage to the genes that control for cell growth.

Cancer registry: An organisation that collects comprehensive information on all new cases of cancer occurring in a defined population.

Cancer site: The body place that a cancer originates in, e.g. lung, breast or prostate.

Cell type: Classification of a cancer according to the type of cell that the tumour resembles. The most common categories include: carcinoma, lymphoma, leukaemia, sarcoma and glioma. Carcinomas represent the most common cancers with sub categories frequently used including adenocarcinoma, squamous cell carcinoma and basal cell carcinoma.

Tensor date: The date at which a patient was last known to be alive or dead.

Census output area: The smallest geographic area commonly used in Northern Ireland. Census output areas are derived from the results of the 2001 Census and aggregate exactly to electoral ward and district council. There are 5,022 COAs in Northern Ireland with an average of 335 persons resident in each area.

Chemotherapy: Treatment of cancer through the use of drugs to kill cancer cells.

Confidence interval: The range of values calculated to have a specified (usually 95%) probability of containing the true value of an observation. Thus the 95% confidence interval for a rate is the range of values within which there is a 95% probability of finding the true value for the rate.

Cohort analysis: The traditional method of survival analysis that is based upon the survival experience of patients diagnosed with cancer during a particular period of time and who have been followed up until the most recent date possible. See observed and relative survival.

Conditional survival: The probability of survival given that survival for a certain length of time has already occurred.

County: A geographic area used in Ireland. There are a total of 32 counties, 26 of which are in Republic of Ireland, the boundaries of which are used for administrative purposes. Local Government in Northern Ireland no longer uses the six counties in Northern Ireland.

Crude rate: The rate per 100,000 persons that an event occurs among a given population. It is calculated by using the formula:

$$C = \frac{R}{N} \times 100,000$$

where C is the crude rate, R is the number of events and N is the population within which the events occur.

Cumulative risk: The risk of an individual developing cancer before age 75 assuming the absence of other causes of death. It is derived using the formula:

$$CR_{0-74} = 100 \left[1 - \exp \left(- \frac{1}{100} \sum_{i=1}^A \frac{a_i t_i}{100000} \right) \right]$$

where a_i is the age-specific rate per 100,000 persons for age class i , t_i is the duration of age class i , A is the number of age intervals between 0 and 74 and CR_{0-74} is the cumulative risk of developing cancer before the age of 75.

Deprivation quintile: The division of census output areas (in Northern Ireland) and electoral districts (in Republic of Ireland) into five groups of approximately equal population size based upon their level of economic deprivation.

Diagnosis: The process whereby the nature of a patient's illness is identified through medical examination.

Direct method: See age-standardised rate.

District council: A geographic area in Northern Ireland defined for Local Government purposes. There are currently 26 district councils in Northern Ireland. District councils are also referred to as Local Government Districts (LGDs).

Ederer II method: See expected survival.

Electoral district: A small geographic area used in the Republic of Ireland.

European standard population: A standard population using the age distribution per 100,000 persons given in the table below. The same age distribution is used for males and females.

Age class	Population	Age class	Population	Age class	Population	Age class	Population
0-4	8,000	25-29	7,000	50-54	7,000	75-79	2,000
5-9	7,000	30-34	7,000	55-59	6,000	80-84	1,000
10-14	7,000	35-39	7,000	60-64	5,000	85+	1,000
15-19	7,000	40-44	7,000	65-69	4,000		
20-24	7,000	45-49	7,000	70-74	3,000	Total	100,000

Excess hazard ratio: The ratio of excess mortality among a group of patients relative to that of a control or baseline group.

Excess mortality: The additional deaths found among a group of patients with cancer, having allowed for the expected mortality rate among persons of the same age and sex in the general population.

Expected survival: The survival expected in a subset of the general population whose characteristics are the same as that of the group of cancer patients being studied. The method used in this report is the Ederer II method, which uses the formula:

$$E_i = \prod_{k=1}^i 1 - \sum_{h=1}^{n_k} \frac{P_k(h)}{n_k}$$

where E_i is the expected survival for a time i after the date of diagnosis, k is the same predefined time interval between the date of diagnosis and i as used in the calculation of observed survival, n_k is the number of patients alive entering interval k and $P_k(h)$ is the probability of a similar person, h , in the general population surviving to the end of interval k . This latter value is taken from life tables derived from population data and deaths from all causes.

Hormone therapy: The treatment of cancer through the addition, removal or blockage of hormones.

ICD10: The tenth edition of the International Classification of Diseases and Related Health Problems, which is published by the World Health Organisation (WHO). It provides a detailed description of known diseases and injuries and is used in the production of morbidity and mortality statistics.

Incidence: The number of new cases of a cancer diagnosed in a particular period for a particular population.

In situ tumour: An early cancer that has not spread to neighbouring tissue.

Kaplein-Meier method: See observed survival.

Lead-time bias: An artificial increase in survival time as measured from the date of diagnosis where earlier detection has not resulted in a delay to the patient's death. The only impact is that patients and services are aware they have cancer for a longer period of time.

Life table: A table that shows the life expectancy of a person at each age and sex. Also usually included in life tables is:

- the probability that a person of a given age will die before their next birthday;
- the number of people out of 100,000 live births who survive to a given age;
- the number of people who die at a given age.

Local Government District: See district council.

Logistic regression: A form of regression used to determine the relationship between variables and a binary (i.e. coded to 0 or 1) dependent variable.

Log-linear model: A mathematical model in which a continuous variable, y , is related to an explanatory variable, x , by the following equation:

$$\ln(y) = mx + b$$

where b is a constant value and m is the gradient of the straight line that best fits the data.

Malignant tumour: A cancerous tumour that can invade and destroy nearby tissue and spread to other parts of the body.

Microscopic verification: A diagnosis of cancer based upon microscopic verification of a tissue specimen including histological confirmation, examination of cytology specimens, and diagnoses of leukaemia based on haematological examination.

Mid-year population estimate: An estimate of the population in a region. Population estimates are based upon the number of births, deaths and migration flows for regions that have occurred since the last population census.

Morphology: The type of cell affected by cancer.

Mortality: The number of deaths from a particular cause for a particular period of time and population.

Mortality:Incidence ratio: The ratio of the number of deaths due to cancer in a given time period to the number of newly diagnosed cases of cancer.

Observed survival: The probability, S_i , that a patient with cancer will be alive at the end of a particular length of time, i , after the date of diagnosis. It is calculated using the formula

$$S_i = \prod_{k=1}^i \left(1 - \frac{d_k}{n_k - \frac{1}{2}w_k} \right)$$

where k is a predefined time interval between the date of diagnosis and i , d_k is the number of deaths from any cause occurring during interval k , n_k is the number of patients alive entering interval k and w_k is the number of patients withdrawn alive during the k^{th} interval.

Odds ratio: The ratio of the odds of an event occurring in one patient group to the odds of it occurring in a baseline patient group.

Passive follow up: A method of cancer registration in which the status of a patient is identified by the matching of cancer registrations with death registrations. This approach is used by both NICR and NCRI.

Pathology: The identification of cancer through the study of cells through a microscope.

Period analysis: An approach used to estimate patient survival for more recent periods of time that cannot be obtained using cohort analysis due to insufficient follow up time. This approach utilises the survival experience of patients still alive in the period of interest rather than of those diagnosed during the period.

Poisson regression: A form of regression that models count data using a log-linear model.

Prevalence: The number of current cases of a disease within a population.

P-value: The probability of an event occurring given a null hypothesis is true. In any statistical tests in this report the null hypothesis is taken to be that there is no difference between two mean values or rates. A small p-value (typically less than 0.05) suggests that the two means or rates tested are significantly different. In this case the result is called statistically significant.

Radiotherapy: The application of radiation to either destroy or reduce the size of malignant tumours.

Relative risk: The ratio of the probability of an event occurring in a group of patients compared to the control or baseline group of patients.

Relative survival: The ratio of the observed survival of a given group of patients to the expected survival for a group of persons in the general population with the same characteristics (usually sex and age, but also country in this report).

Screening: A method of checking for the presence of cancer when there are no signs or symptoms.

Stage: A measure of how far a malignancy has spread in the body. Staging is carried out using a number of laboratory and clinical tests at diagnosis. The most common classification used is the TNM stage that includes information on the extent of the primary tumour (T), the absence or presence of lymph node metastasis (N) and the absence or presence of distant metastasis (M).

Standardised incidence ratio (SIR): The ratio of the number of newly diagnosed cancers observed in a given population to the number of cases expected in a reference population of the same size. The expected number of incidence is calculated by applying a standard set of age-specific rates to the given population. The formula for the standardised incidence ratio (SIR) is:

$$SIR = \frac{\sum_{i=1}^A r_i}{\sum_{i=1}^A \frac{a_i n_i}{100000}}$$

where a_i is the age-specific rate in the reference population, n_i is the observed population in age class i and r_i is the observed number of cases in age class i .

Standardised mortality ratio (SMR): The ratio of the number of cancer deaths observed in a given population to the number of deaths expected in a reference population of the same size. The SMR is calculated in the same manner as the standardised incidence ratio using deaths due to cancer instead of the number of newly diagnosed cases.

Statistical significance: See p-value.

Standardised rate ratio: The ratio of two age-standardised rates which have used the same standard population.

Surgery: An operational procedure conducted to remove cancerous tissue or control its spread. Investigative surgery conducted to diagnose or investigate the presence of cancer is not included in the definition of surgery used throughout this report.

Survival curve: A plot of survival probability against time.

TNM stage: See stage.

Topography: The site/part of the body in which the tumour is present.

Tumour: An abnormal mass of tissue resulting from uncontrolled cell growth and causing a swelling of the body. Tumours may have one of four behaviours: benign, in situ, uncertain or malignant.

Uncertain tumour: A tumour, which at the time of diagnosis, cannot be classified as either benign or malignant.

Vital status: Whether or not a patient is alive or dead at the censor date.

World standard population: A standard population using the age distribution per 100,000 persons given in the table below: The same age distribution is used for males and females.

Age class	Population	Age class	Population	Age class	Population	Age class	Population
0-4	12,000	25-29	8,000	50-54	5,000	75-79	1,000
5-9	10,000	30-34	6,000	55-59	4,000	80-84	500
10-14	9,000	35-39	6,000	60-64	4,000	85+	500
15-19	9,000	40-44	6,000	65-69	3,000		
20-24	8,000	45-49	6,000	70-74	2,000	Total	100,000