Cancer Registry of Norway

Cancer in Norway 2023

Technical Supplement: Statistical Methods

Department of Registration
Section for Analysis and Research

NIPH
Norwegian Institute of Public Health

Cancer in Norway 2023

Technical Supplement: Statistical Methods

Cancer Registry of Norway
Department of Registration

27th June 2024

1 Statistical methods

1.1 Target readership

The target readership for this technical supplement include statisticians and cancer epidemiologists.

1.2 Incidence and mortality rates

Rates are used to measure the frequency at which an event occurs in a defined population during a defined time period. Rates facilitate comparisons across groups of different sizes. Let d and Y denote the number of events and the number of person-years in the population, respectively. In Cancer in Norway 2023 (CIN2023) the mid-year population Y is calculated as the arithmetic mean of the population at the start and end of each calendar year. If the interest lies in calculating a rate over a period of more than one year, we first calculate the annual mid-year population, and then aggregate to determine the total number of person-years.

Rates are reported both as age-specific rates and age-standardised rates (ASR) per 100,000 person-years. The population is divided into 18 distinct five-year age groups. Let d_i and Y_i denote the number of events and the total number of person-years, respectively, for age group i. The age-specific rate r_i per 100,000 person-years, for age group i is then given by

$$r_i = \frac{d_i}{Y_i} \cdot 10^5 \tag{1}$$

The age-standardised rate is calculated as

$$ASR = \frac{\sum_{i=1}^{18} w_i r_i}{\sum_{i=1}^{18} w_i} \tag{2}$$

where w_i is a weight, for age group i, given by the reference population.

Typically the World Standard Population has been used (Doll & al, 1966). CIN2023 uses the age distribution of the Norwegian 2014 mid-year population as the standard population. The population weights for the World Standard Population and the Norwegian 2014 population are given in the table below. One should

note that the World Standard Population places a higher weight on younger age groups and a lower weight on older age groups compared to the 2014 Norwegian population.

i	Age	World (Doll/Segi)	Norway (2014)
1	0–4	12,000	6,039
2	5–9	10,000	6,102
3	10–14	9,000	5,993
4	15–19	9,000	6,349
5	20–24	8,000	6,681
6	25–29	8,000	6,770
7	30–34	6,000	6,690
8	35–39	6,000	6,670
9	40–44	6,000	7,296
10	45–49	6,000	7,207
11	50–54	5,000	6,492
12	55–59	4,000	6,108
13	60–64	4,000	5,575
14	65–69	3,000	5,369
15	70–74	2,000	3,702
16	75–79	1,000	2,663
17	80–84	500	2,063
18	85+	500	2,231
Sum		100,000	100,000

1.3 Cumulative risk

Cumulative risk (CR) is an estimate of the risk of developing a certain type of cancer by a given age. In CIN2023 CR is estimated up to the age of 74. It is defined as

$$CR = 1 - e^{-\left(5 \sum_{i=1}^{N} r_i\right)} \tag{3}$$

where i denotes age group and N denotes the number of age groups used.

1.4 Prevalence

Prevalence is calculated as the number of people in the population that are alive and have been diagnosed with the cancer of interest at some point during their

lifetime.

1.5 Relative survival (net survival)

Estimation was restricted to patients younger than 90 year at diagnosis, and the number of patients in the group to be analysed should be minimum 30.

To estimate net survival the Pohar Perme estimator (Perme & al, 2012) was used applying the stnet command (Coviello & al, 2015) for life-table estimation of net survival, with monthly life-table intervals. National general population life table by sex , year and age (annual) was used in calculation of expected survival. Age standardisation was performed using the Brenner method (Brenner & al, 2004) applying individal weights (Rutherford & al, 2020). The weights used were based on the age distribution of the group to be analysed in the last five-year period, 2019–2023. For all sites combined, sex specific weights were used, while for all other sites, weights were based on the combined male and female age distribution.

From five to three age groups were used for defining weights. The following preference rules were used to derive weights depending on data:

1. Weights based on age distribution quintiles if minimum three patients at start follow-up in each age group.
2. Weights based on age distribution quartiles if minimum three patients at start follow-up in each age group.
3. Weights based on age distribution tertiles if minimum three patients at start follow-up in each age group.

Estimation was not done if failing the last criterion above, or the group to be analysed had less than 30 patients at start follow-up.

When plotting estimates without confidence intervals:

1. Estimates based on < 20 patients were not plotted.
2. For the age group 80–89 estimates were plotted up to 10 years only.

All survival analyses were performed using Stata/MP 18.0 for Windows (64-bit x86-64) Revision 04 April 2024 with stnet.ado version 1.0.8 11Jun2020.

1.5.1 Trends in relative survival

Estimating trends in relative survival were done following the rules outlined above. To increase robustness each point estimate were calculated using survival experience from a larger time period. For years of diagnosis where we have complete five year follow-up (up to and including 2018) the cohort approach was used. That is, the five-year relative survival for 2018 was estimated using patients diagnosed in the period 2014–2018. The five-year estimate for 2017 was estimated using patients diagnosed 2013–2017.

Five-year relative survival for the last year (2023) was estimated using the period-approach (Brenner and Rachet, 2004). The principle of this approach is to fix an observation window, we used 2019–2023, and use all survival experience that occurs within this window in the estimation. This means that patients are left-truncated at the start of the window and right-censored at the end of the window. For five-year relative survival patients diagnosed back to 2014 are included in the analyses, but only with their survival experience occuring from January 1st 2019 and up to five years after diagnosis.

For the years 2019–2022 we use a combination of the cohort and period approaches, a *hybrid approach*. This is implemented by defining a time window, similar to the observation window described for the period approach. Patients diagnosed within the window contribute with all available survival time. Patients diagnosed before the start of the time window have their survival time left-truncated at the start of the time window. Patients diagnosed after the end of the time window are not included in the analyses. For example, the five-year relative survival estimate for 2019 is, thus, based on the survival experience from all patients diagnosed in the time window 2015–2019. In addition, all survival experience occuring from January 1st 2015 for patients diagnosed 2010–2014 are included. The idea behind the hybrid approach is that the incomplete survival for patients diagnosed within the time window, in the example this means patients diagnosed in 2019, are compensated by borrowing survival experience from patients diagnosed before the start of the time window.

1.6 Completeness

Completeness was estimated by the use of the capture-recapture method described in Parkin and Bray (Parkin and Bray, 2009).

This method has been used to estimate the size of a population, and is widely used in field biology to estimate the size of a closed animal population. In that purpose, and briefly explained, animals are captured, marked and released, followed by a new catch (recapture). The number of captured animals in the first catch, the number of new animals in the second catch and the number of recaptured animales are used to estimate the number of uncatched animals.

When this method is used to estimate completeness in a cancer registry context, we assume that cases are registered by two different data sources. Cases registered on pathology reports and/or death certificates (source A) is the first 'catch', and cases registered on clinical notifications (source B) is the second 'catch'. The number of cases registered in source A and/or B is illustrated in the table below.

		Source B	
		Yes	No
Source A	Yes	n_{11}	n_{10}
	No	n_{01}	n_{00}

The completeness (Comp) is then calculated by the following formula:

$$\widehat{Comp} = \frac{n_{11} + n_{10} + n_{01}}{(n_{10} + n_{11}) \cdot \frac{(n_{01}+n_{11})}{n_{11}}}$$

2 References

Doll R, Payne P, Waterhouse J. (Eds) Cancer Incidence in Five Continents: A Technical Report. Springer-Verlag (for UICC), Berlin, 1966.

Perme P, Stare M, Estève J. On estimation in relative survival, Biometrics 2012 Mar; 68(1):113-20.

Coviello E, Seppä K, Dickman P W, Pokhrel A, Estimating net survival using a life-table approach, 2015, The Stata Journal, StataCorp LP, vol. 15(1), pages 173-185.

Brenner H, Arndt V, Gefeller O, Hakulinen T. An alternative approach to age adjustment of cancer survival rates Eur. J. Cancer, 40 (15) 2004, pp. 2317-2322

Brenner H & Rachet B.Hybrid analysis for up-to-date long-term survival rates in cancer registries with delayed recording of incidence cases. Eur. J. Cancer, 40 2004, pp 2494–2501.

Rutherford M J, Dickman P W, Coviello E, Lambert P C. Estimation of age-standardized net survival, even when age-specific data are sparse. Cancer Epidemiology 2020;67:101745.

StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC.

Statistical Methods in Cancer Research. Volume IV - Descriptive Epidemiology IARC Scientific Publications No. 128 Edited by Jacques Esteve, Ellen Benhamou, Luc Raymond. Lyon, 1994. http://publications.iarc.fr/_publications/media/download/3552/d7b5950a320bdfab86b18a7dd79678d88e9bbde2.pdf

Statistics Norway, Statistics Norway provides statistics on population size each year, www.ssb.no/befolkning/statistikker/folkemengde.

Return Address:
Kreftregisteret
P.O. box 5313 Majorstuen
N-0304 Oslo
Norway

Cancer Registry of Norway
Norwegian Institute of Public Health

Postal Address:
P.O. box 5313 Majorstuen
N-0304 Oslo
Norway

Office Address:
Ullernchausséen 64, Oslo

Telephone: +47 22 45 13 00

E-mail: kreftregisteret@kreftregisteret.no
Internet: www.kreftregisteret.no

https://www.kreftregisteret.no/globalassets/cancer-in-norway/2023/cancer-in-norway-2023.pdf
https://www.kreftregisteret.no/globalassets/cancer-in-norway/2023/cancer-in-norway-2023-technical-supplement.pdf